University of Wales, Bangor
College of Natural Sciences

School of Ocean Sciences

PhD thesis presented by Delphine LALLIAS

Genetic linkage mapping in the blue mussel *Mytilus edulis* and the European flat oyster *Ostrea edulis*, and the search for Quantitative Trait Loci of resistance to a disease in *O. edulis*.

Abstract

Genetic linkage maps consist of ordering molecular markers across the genome and require a high number of markers for a good coverage of the genome. Such maps represent a framework which enables the identification and localisation of Quantitative Trait Loci (QTLs) for traits of interest, such as growth or disease resistance, with the final aim of achieving genetic improvement through marker-assisted selection (MAS). Data on bivalves are scarce. No genetic map has yet been constructed in any mussel or flat oyster species.

We report construction of a preliminary genetic linkage map in the blue mussel, *Mytilus edulis* (n=14). Amplified Fragment Length Polymorphism (AFLP) markers were used in a mapping family containing 86 full-sibs issued from a controlled pair mating, applying a double pseudo-test cross strategy. Sex-specific linkage maps consisting of 14 linkage groups were built with MapMaker 3.0 software. In addition, a consensus map was built for 9 homologous pairs based on multiple and parallel linkages of 3:1 markers (segregating through both parents) with JoinMap 4.0 software.

Moreover, the first genetic linkage map for the European flat oyster *Ostrea edulis* (n=10) is presented. AFLP markers and twenty microsatellites were genotyped in a three-generation pedigree mapping family comprising the 2 grand-parents, 2 parents and 92 progeny. Sex-specific linkage maps were built with CriMap software, achieving an estimated genome coverage of 82-84%. Eight linkage groups that were probably homologous between the two parents were identified by the mapping of microsatellites and 3:1 AFLPs. Distorted markers were not randomly distributed across the genome and tended to cluster in a few linkage groups.

Finally, we report the finding of several potential QTLs of resistance and susceptibility to bonamiasis (parasitosis due to *Bonamia ostreae*) in *O. edulis*. The experimental set up consisted of a 6-month trial challenge experiment by cohabitation of wild oysters (overinfected with *B. ostreae*) and tested oysters (two three-generation pedigree segregating families). The disease was transmitted from the wild oysters to the tested oysters, with the first mortalities in the tested oysters occurring after four months of cohabitation. There was a good concordance in the QTLs obtained with three different methodologies used: multi-stage testing strategy, genetic mapping and QTL mapping with QTL express software. The results, even if preliminary, represent a first step towards MAS in the flat oyster.

Contents

List of tables

List of figures

Valorisation of my PhD work

- Publications

Lallias D., Lapègue S., Hecquet C., Boudry P. and Beaumont A.R. AFLP-based genetic linkage maps of the blue mussel *Mytilus edulis*. Animal Genetics 38: 340-349.

Lallias D., Beaumont A.R., Haley C.S., Boudry P., Heurtebise S. and Lapègue S. A first genetic linkage map of the European flat oyster *Ostrea edulis* (L.) based on AFLP and microsatellite markers. In press in Animal Genetics.

- Oral presentations

Lallias D., Lapègue S., Haley C., Heurtebise S., Boudry P. and Beaumont A.R. A preliminary genetic linkage map for the European flat oyster *Ostrea edulis*. Talk given by Sylvie Lapègue at the International Conference and exposition of the world aquaculture society. May 9-13, 2006, Florence, Italy.

Lallias D., Lapègue S., Haley C., Heurtebise S., Boudry P. and Beaumont A.R. A first genetic linkage map for the European flat oyster *Ostrea edulis* L. Talk given by Pierre Boudry at the International Triennal Meeting of the World Aquaculture Society, National Shellfisheries Association and American Fisheries Society-Fish Culture Section. February 27-March 2, 2007, San Antonio, Texas, USA.

Lallias D., Lapègue S., Haley C., Heurtebise S., Boudry P. and Beaumont A.R. A first genetic linkage map for the European flat oyster *Ostrea edulis*- towards the identification of *Bonamia* resistance QTLs. Talk given by myself at the 8th International Marine Biotechnology Conference. March 11-16, 2007, Eilat, Israel.

- Poster presentations

Lallias D., Boudry P., Heurtebise S., Lapègue S. and Beaumont A.R. Establishment of a genetic map in the European flat oyster *Ostrea edulis* L. Poster presented by myself at the Plant and Animal Genome XIV Conference. January 14-18, 2006, San Diego, California, USA.

Lallias D., Lapègue S., Haley C., Heurtebise S., Boudry P. and Beaumont A.R. A preliminary genetic linkage map for the European flat oyster *Ostrea edulis* L. Poster presented by myself at the International Symposium on Genetics in Aquaculture IX. June 25-30, 2006, Montpellier, France.

Lallias D., Hecquet C., Boudry P., Lapègue S. and Beaumont A.R. A first genetic linkage map for the blue mussel *Mytilus edulis*. Poster presented by Pierre Boudry at the International Triennal Meeting of the World Aquaculture Society, National Shellfisheries Association and American Fisheries Society-Fish Culture Section. February 27-March 2, 2007, San Antonio, Texas, USA.

Lallias D., Hecquet C., Boudry P., Lapègue S. and Beaumont A.R. A first genetic linkage map for the blue mussel *Mytilus edulis*. Poster presented by myself at the 8th International Marine Biotechnology Conference. March 11-16, 2007, Eilat, Israel.

Acknowledgements

It has been three years since I started my PhD project. Three years during which I had the opportunity to meet different people, from various fields and various origins. So many people that I want to thank for supporting me and helping me to complete my project. Each contributing in her or his own way. Three years spent between two countries, France (my native country) and Wales (my adopting country).

I would like to thank Philippe Goulletquer, head of the Ifremer laboratory of La Tremblade, and Colin Jago, head of School of Ocean Sciences, for welcoming me into these two research institutes.

All my thanks and gratitude go first to my three supervisors, Andy Beaumont, Sylvie Lapègue and Pierre Boudry, for their support, confiance and encouragements and for giving me the opportunity to start this adventure. I consider myself very lucky to have been part of this European project, AAAG (INTERREG III), even if sometimes I had the impression to be the "black sheep" of this group of researchers working on sustainable aquaculture in various fields other than genetics… It was a good experience for me to attend these AAAG meetings (always in very nice places!) they gave me the opportunity to develop my pedagogical skills in front of a non-geneticist audience!

Andy, thank you for helping me to settle into this new country, to integrate into the small (but friendly!) team of geneticists in SOS and for trusting me before even knowing me! Your advice has been very useful and your attention precious. Thank you for sharing your experience with me, reassuring me and encouraging me during the early hard times where the production of the mussel families was not successful. I remember you telling me that I had three years to complete my PhD and that it was normal not to have any results or successful experiments during the early stages of the project! In other words, I learnt from you patience and perseverance, two fundamental qualities of a researcher. Finally, I would like to thank you for taking the time to read my manuscript, for improving my English and for your constructive comments. I know you were (and are) always very busy, nevertheless you took the time, even your free time, during your week-ends, so thanks for your concern!

Sylvie, it was really nice to share the office with you during 2 years and to create a tropical environment in La Tremblade during the winter!! Thank you for sharing more than science, "toutes ces petites choses de la vie", that are so important!! Also, I would like to thank you for all your invaluable support during this labour-intensive project. When I needed it, you were always there to cheer me up or to help me find a solution or give me advice. I really appreciated the "feminine touch" and "feminine sensitivity" that you brought. I would like to thank you for the useful contacts you gave me so that I can increase my knowledge in genetic and

QTLmapping by meeting researchers from other institutes working in this field (Chris Haley, Christophe Plomion, Catherine Bodénès). These contacts have proved to be very useful and critical for the realisation of my project. Thank you as well for your dedication in reading my thesis, you were always the first to hand back your comments (is it linked to the efficiency of women??).

I would like to thank you, Pierre, for the fruitful discussions about the orientation of this project, for following the advancement of my work despite your "overbooked" timetable! Thank you for taking the time to read part of my thesis and for reviewing the papers.

My thanks should also go to three researchers who advised me and shared with me their knowledge and experience in genetic and QTL mapping.
I had the fantastic opportunity to work with Chris Haley during one month, at the Roslin Institute in Scotland, thanks to a one-month fellowship awarded by the European project "marine genomics". Chris welcomed me into his team, and introduced me to the joys of the Crimap software. During that month, I followed some Msc lectures and tutorials on genetic and QTL mapping at the University of Edinburgh. This training was a critical step of my PhD work, in which I learnt so much! After that experience, I exchanged numerous e-mails with Chris on the type of analyses to perform. I was very impressed by his knowledge and availability, he was always very prompt to reply, and full of advice! At the end of my PhD work, Chris advised me on the best way to analyse my QTL data with the QTL express software and he took the time to explain to me the various analyses done by the different modules of the software.
I would like also to thank Christophe Plomion and Catherine Bodénès from INRA for their advice on the Mapmaker software and on genetic mapping in general. Meeting them has been very interesting and useful for completing my analyses on the establishment of a genetic linkage map in the blue mussel.

All my thanks go to Richard Braithwaite ("Rich") for helping me during the production of the segregating families of mussels in SOS. Rich first initiated me to the pleasure of mussel spawning (going to the lab at 7am to inject KCl into the mantle cavity, putting the mussels into individual jars with sea water and then wait… sometimes up to 10 hours before to collect the gametes!!). He set up the settlement system (downwelling, in a close system with filtered sea water). Without the help of Rich, I don't think I would have been able to produce the segregating families of mussels. Sharing the office with Rich for a few months has been a pleasure. Thank you for your kindness and enthusiasm!

All my gratitude goes to Serge Heurtebise from Ifremer for his crucial role in the production of the different generations of flat oyster families. Serge performed the productions in 2003-2004 and supervised me for the production of 2005. His experience in the production of flat oyster families (from spawning and larval rearing to settlement and growth) was

invaluable. It was a pleasure to work with him, "toujours dans la bonne humeur". Thank you, Serge, for your kindness and the help you gave me in various fields (from hatchery to molecular biology, informatics included!). Merci pour tout!

I wish to thank all the team of pathologists of Ifremer La Tremblade, Isabelle Arzul, Maeva Robert, Sylvie Ferrand, Bruno Chollet and Benjamin Morga, for their involvment, enthousiasm and dedication during the Bonamia challenge experiment. Rien n'aurait été possible sans vous! Votre aide m'a été precieuse.

Special thanks to Christopher Sauvage, PhD student at Ifremer La Tremblade who is working on a similar project. During all the time I spent in La Tremblade, we worked alongside each other, and exchanged a lot about our respective projects. I found in Christopher my closest collaborator and a very good friend. I miss our coffee breaks that we spent chatting and discussing. J'espère que nos routes se re-croiseront un jour! Et je suis sûre que nous resterons en contact!

My thanks should also go to the technical team of the hatchery of Ifremer for welcoming me among them, integrating me into their team and participating in the Bonamia challenge experiment by checking the dead oysters during the week-ends: Raphaël Brizard, Christophe Ledu, Jean-Christophe Billy, Pascal Phelipot, Pascal Schwartel and Florence Cornette in particular. Merci, Flo, pour ton dynamisme au sein de l'équipe! J'étais également ravie de pouvoir plonger avec toi et de partager cette passion! Et merci également pour toute l'aide que tu m'as apportée pendant ce projet!

I would like also to thank Céline Hecquet, my first trainee (and so far the only one). Once more, I really enjoyed this experience that has helped me to develop my pedagogical skills. Despite the long hours, you always kept your enthusiasm and motivation. It was a great pleasure to work with you and to share more than science. Merci pour ta persévérance. J'espère que tu garderas un bon souvenir de cette expérience et qu'elle t'a été profitable pour ta carrière!

Thank you to Tim Croucher and Shelagh Malham for helping me during the first stage of development of the AFLP methodology in SOS. It was a pleasure to share the laboratory bench with them and to exchange our respective experience. I found this experience very stimulating and their advice very useful. Thank you for the good ambience in the lab!

Another friend I met during these 3 years is Johanna. The first time we met, I couldn't believe you were not British! You speak so well, without an accent to my ear! I wish I could speak like you one day! I will try to adapt this French accent that made us laugh so much when I had to say three "r" in the same sentence. So difficult for a native French! Remember "larval rearing room"!! You can be reassured, I still can't say it! Thank you also for being such a good

cook and for making me discover some traditional Swedish food. We spent some really nice evenings sitting around a table, chatting. Tack så mycket!

There are many other people from Ifremer I would like to thank.
Nicole, renommée Victor apres l'inoubliable séance de nettoyage et la réorganisation introduite au sein du laboratoire. Je pense que ça restera gravé dans les annales de l'Ifremer! Merci aussi pour ta bonne humeur et tes fous rires!
Flo, notre bibliothécaire, la parisienne exilée à La Tremblade (tout comme moi).
Véro, notre secrétaire, qui essaie toujours de trouver des solutions et rendre service. Merci pour ta joie et bonne humeur.
Martine, la comptable du labo.
La p'tite Steph, qui n'est (pour l'instant??) pas embauchée définitivement, mais vient faire un séjour à La Tremblade de quelques mois de temps en temps. Merci pour m'avoir aidé à mesurer et peser mes 550 huîtres dans la joie et la bonne humeur. Je n'aurais jamais cru qu'on puisse faire des courses de chariot dans les couloirs de l'Ifremer, et rire autant! Merci pour tes petits cris d'oiseaux et ton rire (communicatif) reconnaissables entre tous!
Karine, pour m'avoir hébergé pendant 5 mois et m'avoir ensuite recommandé auprès du propriétaire pour que je puisse rester dans cette maison si agréable qui n'a connu que des Ifremeriens! Merci pour ta gentillesse et pour les soirées passées à papoter!
Stéphane Robert (le gourmet et bon vivant de la station) pour sa gentillesse et sa bonne humeur. Encore un rire reconnaissable entre tous!
Et tous ceux qui sont passés par La Tremblade ou y sont encore: Maeva, Sylvie, Jean-Côme, Angélique, Yann, Nico, Frederico, Vera, Céline, Christopher, Adeline, Benjamin (qui a pris la relève dans ma petite maison de La Tremblade), Cyrielle, J-C, Christophe pour les agréables soirées passées ensemble- sur la plage, autour d'une bonne table ou d'un barbecue, autour d'une bière à l'Irish pub de Chaillevette (qui l'eut cru!) ou autour d'un verre à Ronce-les-bains.

Special thanks to Matt, the person that should be at the top of my list of thanks for his support and patience during these three years, mainly during the two years spent apart, separated by the English Channel. Thank you for cheering me up each time I needed, thanks for everything you've done for me! Diolch yn fawr!

Mes remerciements pleins d'affection vont bien évidemment à ma famille pour leur support moral et encouragments. Et en particulier à mes parents, qui m'ont rendu visite où que j'aille et quel que soit le temps (Menai Bridge, La Tremblade, Edinburgh, Deganwy). Et bien sûr à ma sœur, mon beau-frère et mon neveu pour être venu me voir au pays de Galles et à La Tremblade! Ces visites m'ont à chaque fois redonné du peps.

My final thanks go to the "Club de badminton trembladais" and the leisure centre of Llandudno junction, as well as all the people that played badminton with me, either occasionally or regularly: Matt, Johanna, Karine, Stephane, Jean-Come, Angélique, Nico, Rich and Kath. Badminton was a fantastic remedy against stress!!

DECLARATION

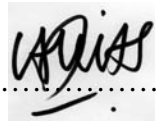
This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed  …….. (Candidate)

Date …16/04/2007……

STATEMENT 1

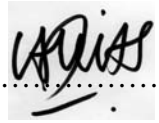
This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Signed  ……. (Candidate)

Date …16/04/2007……

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed  ……. (Candidate)

Date …16/04/2007……

CHAPTER 1

STATE OF THE ART

From quantitative genetics towards Quantitative Trait Loci (QTL) mapping

I- NOTION OF QUANTITATIVE GENETICS

I.1. Quantitative traits or metric characters

Characters where phenotypic variation is continuous and determined by the segregation of a high number of genes are referred to as quantitative traits. Metric characters are measurable, such as anatomical dimensions or physiological functions. Most traits of economical importance in aquaculture are metric, for example growth, survival or disease resistance. The individual loci controlling a quantitative trait are called polygenes or quantitative trait loci (QTL). The frequency distributions of most metric characters approximate more or less closely to normal curves (Falconer and Mackay, 1996). This continuous variation includes the segregation of many genes with small effects.

Quantitative genetics refers to the branch of genetics that covers the study of these quantitative traits. Because of the continuous phenotypic distribution of quantitative traits, classical Mendelian techniques are not applicable to their study. The continuous variation of quantitative traits is instead analysed in terms of means, variances and covariances and its study requires the use of biostatistical methods such as regression analyses or analyses of variances (ANOVA) (Falconer and Mackay, 1996). An important goal of quantitative genetics is to identify the sources of quantitative-trait variation. The determination of the part of the phenotypic variation of the trait which is attributable to genetic or environmental causes relies on the partitioning of the phenotypic variance into components. The relative magnitude of these different sources of variation, estimated by specific experimental designs, is of primary importance to estimate how much of this variation is genetic and therefore heritable and amenable to artificial selection. The proportion of the observed variation that is genetic is known as its heritability (symbol h^2). It is important to note the implications of quantitative genetics in the field of breeding programme. Indeed, quantitative genetics knowledge is required to achieve improvement of a character of economical interest. More specifically, to successfully and efficiently achieve artificial selection, preliminary studies are needed to estimate the potential gain of phenotypic mean expected by selection. Estimates of heritability allow prediction of the response to selection and heritability is therefore a genetic parameter of great importance.

I.2. Variance components and partitioning of phenotypic variance

The study of continuous variation relies on its partition into components attributable to different causes and the estimation of their relative importance. The phenotypic variance V_P can be partitioned into genetic (V_G) and environmental (V_E) components:

$$V_P=V_G+V_E$$

A further step consists of partitioning the genetic component of variance:

$$V_G=V_A+V_D+V_I$$

V_A represents the additive genetic variance, V_D the dominance variance and V_I the epistatic variance. So, the total variance V_P can be written

$$V_P=V_A+V_D+V_I+V_E$$

In conclusion, the genetic component of variance is made up of the additive genetic variance V_A (part of the genetic variance which is heritable and transmitted from parents to offspring) and the non-additive genetic variance (V_D+V_I) which results from interactions among alleles at a locus or to interactions among loci (those associations are broken at meiosis and are therefore not transmitted).

Two ratios are of great importance in quantitative genetics and in breeding programmes. First, the ratio V_G/V_P expresses the part of the phenotypic variation in the character of interest that can be attributed to genetic variation. This ratio includes both the additive and non-additive genetic component. However, only the additive effects can be directly transmitted to the offspring and are useful to explain the degree of resemblance between relatives or to predict the potential response to selection. Therefore the ratio V_A/V_P, based on the only part of the genetic component which is amenable to selection, is of greatest importance. The ratio V_G/V_P is called broad sense heritability, V_A/V_P narrow sense heritability and either type of heritability has the symbol h^2. Heritability ranges from zero (no genetic variance) to 1.00 (no environmental variance).

The term “environment” gives rise to a source of phenotypic variation whose causes are complex and not always well identifiable. This type of variation represents a major component of the total variation for quantitative traits, which frequently exceeds the genetic variation (Kearsey and Pooni, 1998). Therefore, environmental variance is a source of error which reduces the

accuracy of the estimation of genetic parameters. It is very important to develop experimental designs which limit this source of variation.

I.3. Principle of estimation of the components of variances and the heritability.

In practice, quantitative genetic study is based on the development of experimental designs that allow partition of the observed variance of a trait into various components. The analysis is based on the differences of phenotypic values of the trait of interest between relatives of known relationship.

I.3.1. Estimation of heritability: Parent-offspring regression

The regression of offspring on parent is a very useful way of estimating the narrow-sense heritability because of the direct relation between the slope of the regression line and the additive causal component of variance. In the case of the regression of the offspring mean phenotypic value on the values of one parent, the least-square slope b_{OP} of the regression is an estimate of half the heritability:

$$b_{OP} = \frac{1}{2}\frac{V_A}{V_P} = 0.5\ h^2$$

In the case of the regression of the offspring mean phenotypic value on the mid-parent values (mean of the 2 parents), the slope of this regression $b_{O\bar{P}}$ is a direct estimate of the heritability h^2 (Figure 1).

$$b_{O\bar{P}} = \frac{V_A}{V_P} = h^2$$

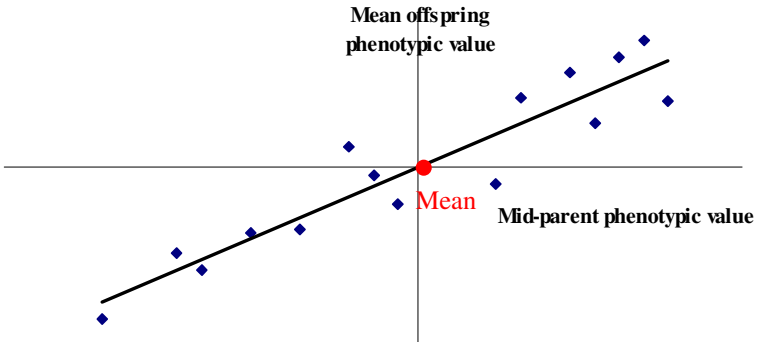


Figure 1. Regression of offspring on mid-parent value. Each point represents an experiment in which the mid-parent value (mean of the two parents values) is plotted against the mean phenotypic value of their offspring. The black line represents the least-square regression, whose slope allows estimation of h^2.

I.3.2 Sib analysis

This approach relies on the analysis of families: half-sib families, full-sib families and combinations of both types of families. Each of these family structures permits one to partition the total phenotypic variance into within- and between-family components. These observed components of variance can be interpreted in terms of covariances between relatives and then be related to the causal components of variance (V_A, V_D, V_E), allowing the estimation of h^2.

- Paternal half-sib design

In this type of mating (Figure 2), each male is mated to different females; one offspring per female is analysed. The offspring have the same father but different mothers and are called half-sibs (HS).

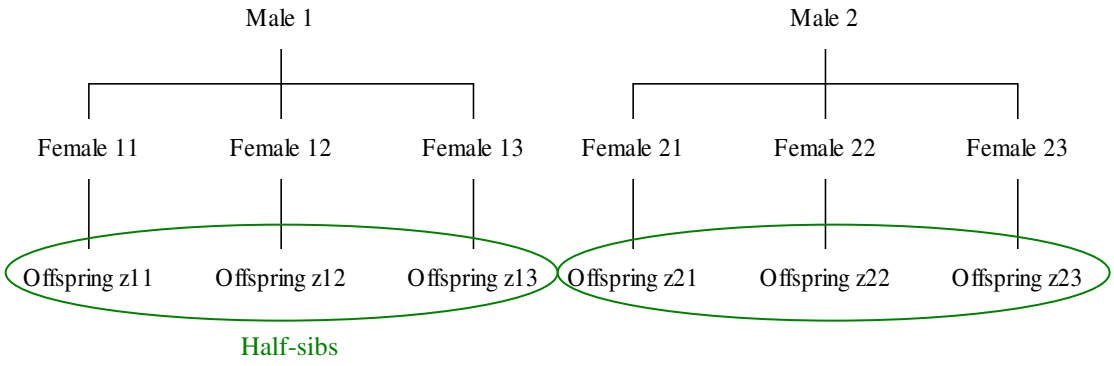


Figure 2. Illustration of a paternal half-sib mating design where individual males are mated to several females.

This experimental design permits estimation of the heritability by using the estimation of the sibs correlations. Indeed,

$$t_{HS} = \frac{COV_{(PHS)}}{\sigma_T^2} = \frac{\sigma_s^2}{\sigma_T^2} = \frac{1}{4}\frac{V_A}{V_P}$$

where $COV_{(PHS)}$ is the paternal half-sib covariance, t_{HS} being the intraclass correlation when using a half-sib mating design, σ^2_T the overall phenotypic variance and σ^2_s the variance between sires. The narrow-sense heritability is then estimated as $4t_{HS}$. Paternal half-sibs analysis has been chosen because this mating design is generally preferred: only the additive genetic effects account for the resemblance between relatives. On the contrary, in the case of maternal HS

design, the estimation of the heritability is inflated by an environmental source of variation due to maternal effects.

- Nested or hierarchical mating design

This design allows the collection of both half-sib and full-sib families (Figure 3).

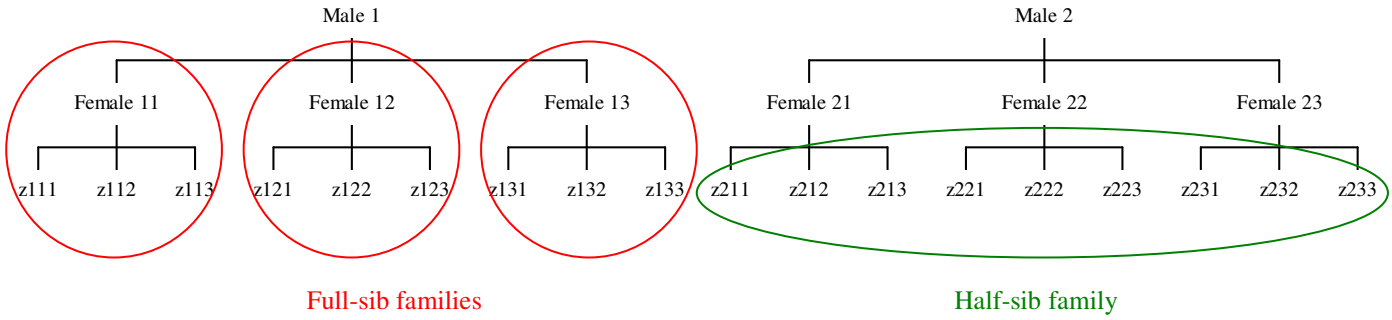


Figure 3. Illustration of a hierarchical mating design, generating half-sib and full-sib families.

In the case of the hierarchical design, the analysis is carried out by a nested ANOVA, as the effect of the female will be studied within the effect of the sire.

$$\sigma_T^2 = \sigma_s^2 + \sigma_d^2 + \sigma_e^2$$

with σ^2_T the overall phenotypic variance, σ^2_s the variance between sires, σ^2_d the variance between dams nested within sires and σ^2_e the within-progeny variance.

The computation of the mean squares (MS) allows estimation of these three components of variance. The covariance between half-sibs is due only to the additive genetic variance: as only one parent is in common, the terms of between-alleles interaction variance (either dominance deviation or epistatic interaction) do not contribute to the resemblance between half-sibs. On the contrary, full sibs have both parents in common: their covariance will have an additive and dominance component (epistatic interactions are often neglected and included in the term V_A). Moreover, full sibs share a common environment, due mostly to maternal effects. The sibs or intraclass correlations can be expressed as:

$$t_{(HS)} = \frac{COV_{(HS)}}{\sigma_T^2} = \frac{\sigma_s^2}{\sigma_T^2} = \frac{1}{4}V_A \qquad \text{and} \qquad t_{(FS)} = \frac{COV_{(FS)}}{\sigma_T^2} = \frac{\sigma_s^2 + \sigma_d^2}{\sigma_T^2}$$

where $t_{(HS)}$ and $t_{(FS)}$ are the half-sib and full-sib intraclass correlations, $COV_{(HS)}$ and $COV_{(FS)}$ are the half-sib and full-sib covariances.

Therefore the nested mating design will allow both the estimation of the additive variance (hence the heritability), based on $t_{(HS)}$ and the non-additive variances.
When the dam component is not significantly different from the sire component, one can conclude that the non-additive genetic source of variation and the maternal effects are not a significant source of variation of the character studied. The narrow sense heritability can then be estimated from either the sire component or from the dam component ($h^2 = 4\frac{\sigma_d^2}{\sigma_T^2}$). A better estimate is to use both the sire and the dam components: $h^2 = 2t_{(FS)}$. Otherwise, when the dam component is significantly greater than the sire component, this experimental design does not allow estimation of the relative contribution of the dominance deviations and the environmental source of variation (maternal effects).

I.3.3 Maximum likelihood statistical procedures

Both approaches previously set out (regression of parent mean and offspring and ANOVA, paternal HS or FS/HS method) allowed estimation of the causal components of variance as simple linear functions of the observed covariance between relatives.

The maximum likelihood procedure (Falconer and Mackay, 1996) represents an alternative for the estimation of the causal components which is now more commonly used. This approach utilizes all the information available, does not require balanced designs and can take into account the selection of parents. Generally speaking, this approach can accommodate any type of data, without any requirement on the properties of the data (ANOVA analysis requires the normality of the data, the independence of the samples and the homogeneity of the variances). An iterative computer algorithm is used to find the combination of the parameters which maximise the likelihood function: the MLEs or Maximum Likelihood Estimates. Various types of procedures can be employed, depending on the data. The most commonly used is the restricted procedure (REML) which adjusts the observations for the estimates of the fixed effects.

II- APPLICATIONS OF QUANTITATIVE GENETICS INTO BREEDING PROGRAMMES IN AQUACULTURE

The estimates of heritability are characteristic of a particular study with its specific conditions. It is therefore not advisable to transpose an estimate from one population to another or for the same population in a different environment. Thus, before implementing a breeding programme, it is advisable to estimate the heritability of the characters of economical importance in situations as close as possible to those in which artificial selection is to be carried out.

Several publications have addressed the importance and potential of genetics in aquaculture, and particularly the role of quantitative genetics (Newkirk, 1980; Wilkins, 1981; Gjedrem, 1983; Beaumont, 1994b; Sheridan, 1997). The development of a breeding programme is driven by the need to improve economically important traits by artificial selection and to reduce the production costs. Quantitative genetic selection may be used in a living organism to improve any trait for which there exists additive genetic variance. Selection can be carried out over many generations, with progress resulting in each, until the genetic variance for the trait is exhausted (Crenshaw et al., 1996). According to Ward et al. (2000), one or more traits of economical interest may be chosen for improvement, but these traits must be variable. Indeed, if there is no variation in the population, selection will be ineffective. Finally, a part of this variation must result from genetic variation, more precisely additive effects (the only one to be directly transmitted to the offspring). The heritability for the character of interest must thus be estimated in the population studied and according to its value, mating schemes and selection methods need to be designed. Generally, heritability values of 0.20 or larger indicate that a genetic gain can be achieved through the application of selective breeding programmes (Newkirk et al., 1977).

Several methods of selection can be undertaken: individual or mass selection, family selection, within-family selection and combined or index selection (Kearsey and Pooni, 1998). Mass selection consists of the selection of individuals based on their phenotypic values. A certain proportion of the base population is selected and interbred to form the next generation. This type of selection relies on the assumption that the phenotypic values are correlated to the breeding values of the individuals (hence, the heritability of the trait must be high). The other types of

selection can be gathered under the term "genotypic selection" because these methods select individuals based on the performance of their relatives (Pillay, 1990). Family selection consists of rearing several families (issued from single pair matings) under identical conditions to determine the ones to be maintained for breeding. Selection is made on the family mean because the families with the highest mean performance will be selected.

Index selection is the most efficient type of selection because it optimises genetic information by allocating appropriate weights to the between- and within-family effects (Kearsey and Pooni, 1998). Mass selection is more efficient than family or within family selection for intermediate to high heritabilities (Kearsey and Pooni, 1998). Therefore, mass selection, due to the simplicity of its realisation, is of particular interest when a single trait is chosen for improvement, especially if its heritability is high. However, applying combined individual and family selection will result in a higher genetic gain and will be more effective for the improvement of traits with low heritabilities (Falconer and Mackay, 1996).

II.1. Estimation of realised heritability in selection experiments

When carrying out mass selection experiments, the calculation of the "realised" heritability is based on the response to selection R and the selection differential S:

$$R = h^2 S$$

The response to selection R is the change of the population mean produced by selection and is therefore the difference between the mean phenotypic value of the offspring issued from the selected parents and the mean phenotypic value of the offspring issued from the whole population before selection. The selection differential S is the average superiority of the selected parents and is the mean phenotypic value of the selected individuals expressed as a deviation from the population mean (Falconer and Mackay, 1996). The realised heritability is then estimated by the ratio R/S (Figure 4).

Heritability normally does not decrease during the first five to ten generations of selection when the initial population size of the broodstock is large (Falconer and Mackay, 1996).

Therefore, the heritability estimate obtained in one generation can be used to predict the gain in the mean value after a certain number generations of selection.

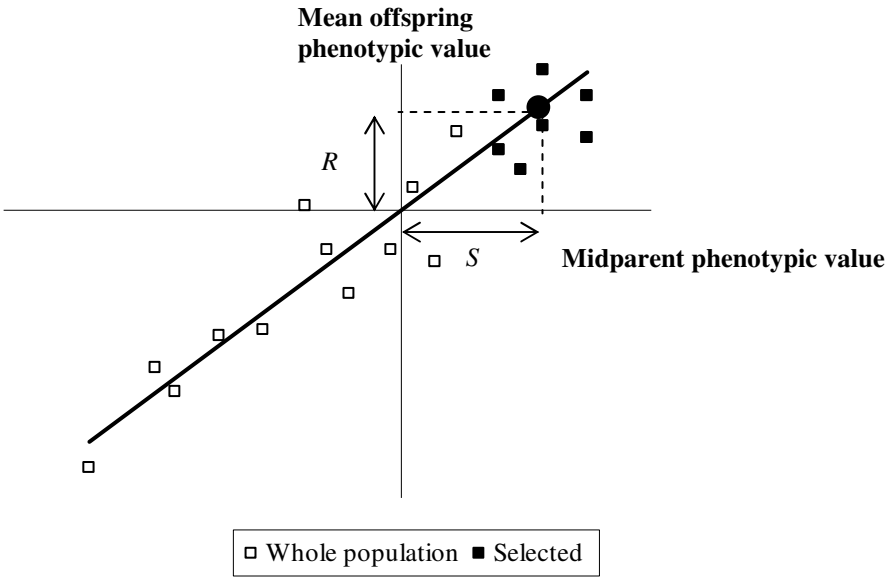


Figure 4. Estimation of realised heritability. The open squares represent the mid-parent values plotted against the mean value of their offspring in the whole population. The closed squares represent the individuals selected. The axes intersect at the mean value of the whole population. The closed circle is the mean of the selected individuals. *R*=response to selection, *S*=selection differential. The realized h^2 is estimated by the ratio *R*/*S*.

Environmental factors (such as climatic, nutritional or management of the cultures) can vary from one generation to the next. Therefore, the comparison between the means of the selected and non selected populations is more efficient when a control population issued from crosses of randomly selected individuals (from parental population) is reared in the same conditions as the offspring of the selected population. The use of divergent selection improves the precision of the heritability estimate (Falconer and Mackay, 1996). In this case, selection of the character is made in both directions, for both an increase and a decrease of the character. Each selected line acts as a control for the other and the response is measured as the divergence between the two lines.

II.2. Review of heritability estimates in various shellfish species

The most commonly studied traits of economical importance in the shellfish industry have been growth and survival. The heritability for these quantitative characters has been studied at the larval, juvenile and adult stages.

One can conclude that the potential for improvement of growth through selection in these species is important. The estimates vary according to the study and the life stage, sometimes associated with a large standard error. However most of the studies report a h^2 value greater than 0.2. Mass selection can therefore be a successful way of improving growth rate (Table 1).

There are fewer estimates of survival (Table 2). Results differ considerably between the studies but some studies affirm that selective breeding programmes could efficiently improve survival (Lannan, 1972; Lannan, 1980a; Jonasson et al., 1999; Degremont et al., 2003). Moreover, non-genetic or maternal effects seem to have an important influence on the variation of the trait (Mallet et al., 1986).

Other traits studied include shell shape and intra-specific competition. Wada (1986) studied the heritability of the shell convexity as an indicator of the shell shape in the Japanese pearl oyster. The responses to selection in the first and second generation as well as in the third generation indicate heritabilities for this trait of 0.32 and 0.35 respectively. The shell shape of the pearl oyster may then be improved through selective breeding. Heritability estimates for shell shape were reported to be 0.31±0.18 and 0.13±0.12 at 12 and 18 months of age respectively in *Crassostrea gigas* (Lannan, 1972). Another interesting study analysed the effect of intra-specific competition on growth in *Mytilus galloprovincialis* (Brichette et al., 2001). These authors demonstrated low heritability estimates (around 0.10) but concluded that significant selection response could be obtained when taking competition effects into account in a breeding programme.

To conclude, it is important to highlight some features common to several studies. First, the potential for improving quantitative traits of economical importance (growth, survival, shape of shell) through a selective breeding plan is real. Secondly, there is a between-generation variation in selection response. According to Newkirk and Haley (1982), the response of selection for weight in *Ostrea edulis* is less in the second generation than in the first one. In addition, Wada (1986) reported a decrease of response in the third generation compared to the first and second one for shell width. This can be due to sampling error, genetic drift, environmental variation or inbreeding depression. A more accurate estimate of the heritability and response to selection is best estimated by averaging the results over several generations (Falconer and Mackay, 1996). Finally, there is a tendency for the heritability estimate to increase with age (Mallet et al., 1986; Toro and Newkirk, 1990; Jonasson et al., 1999; Perez-Rostro and Ibarra, 2003) when the studies are based on HS analysis.

Table 1. Review of heritability estimates for growth in various shellfish species.
FS : full sibs ; HS : half sibs ; PHS : paternal half sibs (one male mated to several females). SR: standardized response; SSD: standardized selection differential.
ANOVA: analysis of variance; REML: restricted maximum likelihood; df-REML: derivative-free restricted maximum likelihood.
h^2_r: realised heritability; h^2_{ns}: narrow-sense heritability; h^2_{bs}: broad-sense heritability. HH, LL, HL, LH: crosses between high (H) and/or low (L) lines.

Species	Character measured	Method of analysis	Heritability estimate	Age of estimation of h^2	References
Ostrea edulis	Weight, length, width (multivariate analysis)	Parent/offspring regression Parents selected at 3 years	$h^2_r \leq 0.10$	1 year	Ruzzante (1986)
	Total live weight	Parent/offspring regression 1st generation of selection High selected lines versus control lines Selection for weight at 2 years 2 experiences at 2 different years	$0.39<h^2_r<0.72$ $0.09<h^2_r<1.18$ Gain in weight: 8 to 38%	2 years	Newkirk and Haley (1982)
	Total live weight	Parent/offspring regression High selected lines versus control lines 2nd generation of selection	$h^2_r=0.12$ not significantly different from 0	2 years	Newkirk and Haley (1983)
	Weight	Parent/offspring regression Divergent selection Selection for live weight at 6 months 18 high, 6 low families	$h^2_r=0.136\pm0.118$	6 months	Toro and Newkirk (1990)
			$h^2_r=0.243\pm0.202$	18 months	
	Length		$h^2_r=0.112\pm0.041$	6 months	
			$h^2_r=0.194\pm0.070$	18 months	
	Weight	Selected populations for resistance to *Bonamia ostreae*	$0.27<h^2<0.84$		Naciri-Graven et al. (2000)
Ostrea chilensis	Live weight	ANOVA included effect of depth, group (high/low) and interaction Divergent selection High and low selected lines for live weights. Simultaneously, control lines. Selection at 40 months	h^2_r high=0.69±0.11 h^2_r low=0.35±0.08	8 months	Toro et al. (1994) Toro et al. (1995)
			h^2_r high=0.55±0.10 h^2_r low=0.24±0.06	14 months	
			h^2_r high=0.43±0.18 h^2_r low=0.29±0.13	27 months	
	Shell length		h^2_r high=0.70±0.10 h^2_r low=0.27±0.06	8 months	
			h^2_r high=0.63±0.09 h^2_r low=0.36±0.07	14 months	
			h^2_r high=0.45±0.12 h^2_r low=0.31±0.11	27 months	

Saccostrea cucullata	Whole weight	Weighted least square analysis Selection at 15 months 3 selected groups (high, medium, low) for weight reared in the same nets	h^2_r=0.277±0.006 (slope of regression of SR on SSD)	Weight recorded every 3 months until 15 months	Jarayabhand and Thavornyutikarn (1995)
Crassostrea virginica	Growth	Hierarchical mating Sib analysis. FS and paternal HS Only one set of experiment	h^2_{ns}=0.24	2 weeks	Longwell and Stiles (1973)
	Growth rate	Selection experiments	h^2_r=0.93	33 days postsetting	
	Log length x width	Factorial mating Sib analysis. FS and HS	0.26<h^2_{ns}<0.50 3 experiments	6 and 16 days	Newkirk et al. (1977)
	Shell length	Factorial mating 9 FS families; maternal and paternal HS Nested ANOVA 3 experiments	h^2_{ns}=0.44±0.21	7 days	Losee (1978)
			h^2_{ns}=0.40±0.20	14 days	
			h^2_{ns}=0.55±0.26	21 days	
			h^2_{ns}=0.50±0.30	6 weeks postsetting	
	Shell width		h^2_{ns}=0.54±0.32	6 weeks postsetting	
	Shell length	HS analysis. 25 families Population: 3 generations of selection	h^2_{ns}=0.44±0.14	7 months	Davis (2000)
	Live weight		h^2_{ns}=0.51±0.15	7 months	
Crassostrea gigas	Size (length, weight, height)	FS analysis (9 matings at 12 months, 11 matings at 18 months). 15 individuals analysed per mating ANOVA	h^2_{bs}=0.93±0.28	12 months	Lannan (1972)
			h^2_{bs}=0.15±0.14	18 months	
	Total weight		h^2_{bs}=0.33±0.19	18 months	
	Shell weight		h^2_{bs}=0.32±0.30	18 months	
	Meat wet weight		h^2_{bs}=0.37±0.20	18 months	
	Meat/total weight		h^2_{bs}=0.46±0.22	18 months	
	Body size	Factorial and hierarchical crosses 192 families. ANOVA	High h^2	Harvest	Hedgecock et al. (1991)
	Yield (wet weight/bag)	FS families (crosses between 9 families selected for high yield; crosses between non selected oysters)	h^2_r=0.42	2 years	Langdon et al. (2003)
	Weight	Nested design. 43 FS families nested within 17 HS families. 3 sites. Nested ANOVA. Significant h^2_{ns} when 3 sites pooled.	h^2_{ns}=0.07±0.07 to 0.15±0.08	6-8 months	Dégremont et al. (2007)

Pinctada fucata martensii (Japanese pearl oyster)	Shell width	First and second generations of response to selection	h^2_r=0.31	3 years	Wada (1984)
	Shell width	Regression parents/offspring Third generation response Base population: selected for yellow coloration in prismatic layer	h^2_r=0.467		Wada (1986)
Mytilus edulis	Length	Nested mating design. 10♂, 30♀ Natural population, not selected Maximum likelihood estimation of the variance components	h^2_{ns}=0.11±0.02	16 days	Mallet et al. (1986)
			h^2_{ns}=0.62±0.06	300 days	
			h^2_{ns}=0.92±0.27	700 days Env.1	
			h^2_{ns}=0.22±0.07	700 days Env.2	
	Length	Nested mating design. 9♂, 27♀	h^2_{ns}=0.8±0.5	14 days	Stromgren and Nielsen (1989)
			h^2_{ns}=0.5±0.3	28 days	
	Length	Shell growth gain of juveniles: 28 to 35%	h^2_{ns}=0.6±0.3	4.5 months	Stromgren and Nielsen (1989)
			h^2_{ns}=0.5±0.2	13 months	
	Growth rate		h^2_{ns}=0.9±0.7	4.5 months	
Mytilus chilensis	Length	Nested mating design. 8♂, 24♀ 8 HS and 24 FS families Nested ANOVA 3 different algal concentrations (low, medium, high). Interaction genotype-food tested by ANOVA	h^2_{ns}=0.5±0.3 (low) h^2_{ns}=0.8±0.3 (high)	6 days	Toro and Paredes (1996)
			h^2_{ns}=0.4±0.3 (low) h^2_{ns}=0.7±0.1 (high)	12 days	
			h^2_{ns}=0.6±0.2 (low) h^2_{ns}=0.9±0.3 (high)	20 days	
Argopecten irradians concentricus Scallop	Shell length	Control line parent. Selected line for high growth.	h^2_r=0.206	10 and 20 weeks	Crenshaw et al. (1991)
Placopecten magellanicus Scallop	Shell length	20 FS families Single pair matings (20♂, 20♀)	h^2_{bs}=1.10±0.17	4 days	Jones et al. (1996)
			h^2_{bs}=1.24±0.40	14 days	
			h^2_{bs}=1.21±0.38	21 days	
Mercenaria mercenaria Clam	Shell length	Nested mating design. 95 FS, 31 PHS Nested ANOVA and REML	h^2_{ns}=0.74±0.07 adjusted for density combined sire+dam	9 months	Rawson and Hilbish (1990)
	Standard length	Control lines. Selected lines. 3 sub-groups spawned at 3 dates.	h^2_r=0.42±0.10	2 years. Expce 1	Hadley et al. (1991)
			h^2_r=0.43±0.06	2 years. Expce 3	
	Shell length	Spawning of 177 selected parents and 177 control parents. Selection at 2 y	h^2_r=0.409 moderate density	2 years	Crenshaw et al. (1996)

Haliotis rufescens Red abalone	Length	Hierarchical mating design 29♂, 88♀. 100 FS and HS families df-REML algorithm	h^2_{ns}=0.08	8 months	Jonasson et al. (1999)
			h^2_{ns}=0.06	10 months	
			h^2_{ns}=0.27	18 months	
			h^2_{ns}=0.34	24 months	
Pennaeus monodon Prawn	Length	Hierarchical design Nested ANOVA h^2 based on sire components	h^2_{ns}=0.08±0.10 ns	6 weeks	Benzie et al. (1997)
			h^2_{ns}=0.12±0.07 ns	10 weeks	
	Wet weight		h^2_{ns}=0.12±0.02	6 weeks	
			h^2_{ns}=0.10±0.002	10 weeks	
Penaeus japonicus Kuruma prawn	Weight	Divergent selection : 13 HH, 13 LL, 5HL and 3 LH families Selection at 6 months for weight Regression offspring/midparent	h^2_r=0.165 (HH) h^2_r=0.315 (LL) h^2_r=0.234 (HH-LL) h^2_{ns}=0.277±0.083	6 months	Hetzel et al. (2000)
Penaeus vannamei Pacific white shrimp	4 size traits	60 FS families REML analysis	h^2_{bs}=0.15-0.22±0.17	17 weeks	Perez-Rostro and Ibarra (2003)
			h^2_{bs}=0.24-0.32±0.17	23 weeks	
			h^2_{bs}=0.28-0.35±0.18	29 weeks	
	Growth rate		h^2_{bs}=0.63-0.84±0.20	Δ17-23 weeks	
			h^2_{bs}=0.34-0.54±0.20	Δ23-29 weeks	
Macrobrachium rosenbergii	Weight	Nested design. 50 FS and HS families	h^2_{bs}=0.35±0.15	Juvenile	Malecha et al. (1984)

Table 2. Review of heritability estimates for survival in various shellfish species. See Table 1 for the abbreviations used.

Species	Character measured	Method of analysis	Heritability estimate	Age of estimation of h^2	References
C. gigas	Survival	FS analysis. ANOVA	h^2_{bs}=0.31±0.06	Larvae	Lannan (1972)
	Survival	Factorial analysis Arcsin transformation ANOVA	h^2_{ns}=0.13 Experience A h^2_{ns}=0.34 Experience B		Lannan (1980b)
	Survival	Nested design. 43 FS families nested within 17 HS families. 3 sites.	h^2_{ns}=0.47±0.20 to 1.08±0.46 h^2_{bs}=0.07±0.04 to 0.38±0.13	6-8 months	Dégremont et al. (2007)
M. edulis	Survival	Nested design. 10♂, 30♀ Natural population, not selected Maximum likelihood estimation of the variance components	h^2_{ns}=0.0; significant non additive or maternal variance	16 days	Mallet (1988)
			h^2_{ns}=0.0 or non significant	700 days	
H. rufescens	Percentage survival	Hierarchical mating design 29♂, 88♀. 100 FS and HS families df-REML algorithm	h^2_{ns}=0.11 (Observed scale) h^2_{ns}=0.33 (underlying liability scale)	4 months	Jonasson et al. (1999)

II.3. Review of heritability estimates in Salmonids: potential for the shellfish industry

As selective breeding programmes are more advanced for finfish than for shellfish, it is interesting to give some examples of the heritability estimates of economically important traits in fishes. On the model of what has been done in fisheries breeding programmes, some traits (such as meat quality or meat colour) could thus be selected in shellfish with important economical repercussions. Naturally, shellfish breeders want to improve the growth rate to attain market size faster and reduce the production costs. However, the quality of the product is a more and more important feature that must not be neglected.

Some examples are cited in Table 3 that have been selected from the extensive literature. Most of them concern Salmonids. For more details, see Gjedrem (1983, 2000) who reviewed the heritability estimates for various economically important traits in rainbow trout, Atlantic and Pacific salmon, carp, catfish and tilapia. The analysis of this table shows that a great variety of economical traits can be improved by selection. Selection for many of these characters could be important in the shellfish industry: meat colour, gonad weight, age at sexual maturity, meat quality or resistance to disease.

Table 3. Review of heritability estimates in Salmonids.

Condition factor=weight x 100/length^3
Family-Tie: sort of factorial mating; each sire mated to 3 females; dams mated to 1 to 2 males
CV: coefficient of variation
h^2_{ns}: narrow-sense heritability; h^2_{bs}: broad-sense heritability
ANOVA: analysis of variance; REML: restricted maximum likelihood; df-REML: derivative-free restricted maximum likelihood

Species	Economical trait	Heritability estimate	Method	References
Atlantic salmon *Salmo salar*	190 days weight	h^2_{ns}=0.08	Nested ANOVA	Refstie and Steine (1978)
	190 days length	h^2_{ns}=0.12		
	Weight ungutted	h^2_{ns}=0.44=±0.11	Nested ANOVA	Gjerde and Gjedrem (1984)
	Weight gutted	h^2_{ns}=0.44±0.11		
	Length	h^2_{ns}=0.35=±0.10		
	Meatiness	h^2_{ns}=0.16=±0.05		
	Meat colour	h^2_{ns}=0.01=±0.03		
	Liver colour	h^2_{ns}=0.04=±0.02		
	Maturity	h^2_{ns}=0.39		
	Survival	h^2_{ns}=0.04±0.04	Nested ANOVA Data adjusted	Jonasson (1993)
	190 days weight	h^2_{ns}=0.16±0.02		
	190 days length	h^2_{ns}=0.10±0.02		
	Body weight	h^2_{ns}=0.36±0.11	Nested mating df-REML	Jonasson et al. (1997)
	Return rate	h^2_{ns}=0.08		
Arctic charr *Salvelinus alpinus*	Length	h^2_{ns}=0.31-0.55±0.15	Mid-parent offspring regression	Nilsson (1990)
	Weight	h^2_{ns}=0.31-0.49±0.14		
	Condition factor	h^2_{ns}=0.07-0.53±0.14		
	Length	h^2_{bs}=0.34-0.55±0.08	FS analysis ANOVA	
	Weight	h^2_{bs}=0.40-0.49±0.10		
	Condition factor	h^2_{bs}=0.32-0.56±0.10		
Chinook salmon	1st winter Weight	h^2_{ns}=0.34±0.07	Family-Tie. df-REML	Winkelman and Peterson (1994)
	1st winter Length	h^2_{ns}=0.32±0.10		
Rainbow trout	140 days length	h^2_{ns}=0.06	Nested ANOVA Data adjusted	Refstie (1980)
	140 days weight	h^2_{ns}=0.20		
	Weight ungutted	h^2_{ns}=0.19=±0.11	Nested ANOVA	Gjerde and Gjedrem (1984)
	Weight gutted	h^2_{ns}=0.19±0.12		
	Length	h^2_{ns}=0.16=±0.11		
	Meatiness	h^2_{ns}=0.14=±0.06		
	Meat colour	h^2_{ns}=0.06=±0.08		
	Liver colour	h^2_{ns}=0.04=±0.05		
	Maturity	h^2_{ns}=0.21		
	Forklength	0.13±0.17 to 0.33±0.20	Factorial design h^2_{ns} estimated at 2.5 and 4 years	McKay et al. (1986)
	Weight	0.17±0.19 to 0.38±0.22		
	Condition factor	0.46±0.24 to 0.66±0.27		
	Age maturity	h^2_{ns}=0.21=±0.14		
	Body shape	h^2_{ns}=0.13-0.25	Nested design REML	Gjerde and Schaeffer (1989)
	Ungutted weight	h^2_{ns}=0.21		
	Gutted weight	h^2_{ns}=0.21		
	Body length	h^2_{ns}=0.18		
	Condition factor	h^2_{ns}=0.19		
	Gonad weight	h^2_{ns}=0.23		

Rainbow trout	Belly thickness	h^2_{ns}=0.40	Nested design REML	Gjerde and Schaeffer (1989)
	Meat colour	h^2_{ns}=0.27		
	Abdominal fat	h^2_{ns}=0.25		
	Ungutted weight	0.28±0.07 to 0.51±0.04	Nested design Multiple trait animal model. Estimation at 1, 1.5, 2 and 2.5 years	Elvingson and Johansson (1993)
	Length	0.41±0.08 to 0.53±0.08		
	Belly thickness	h^2=0.31±0.07		
	Gonad weight	h^2=0.26±0.06		
	Visceral weight	h^2=0.74±0.09		
	Condition factor	0.28±0.07 to 0.51±0.04		
	215 days weight	h^2_{ns}=0.35±0.30	Factorial design REML	Henryon et al. (2002)
	215 days length	h^2_{ns}=0.53±0.27		
	Feed conversion efficiency	CV moderate h^2 not estimated		
	VHS resistance	h^2_{ns}=0.13		

II.4. Estimation of the phenotypic and genetic correlations

Estimating phenotypic and genetic correlations between traits at a particular age or between the same trait at different ages is of primary interest for commercial hatcheries.

It is important to identify and detect the correlated response of the selection of one trait on another trait. Identifying negative genetic correlations between two traits demonstrates that the selection of one trait will act in the wrong direction on another trait, reducing the power of the selection if both traits are of economical value. Toro and Newkirk (1990) reported a very high positive genetic correlation between live weight and shell height in *O. edulis*, suggesting that selection for one trait will lead to a correlated response in the other trait. Toro et al. (1994, 1995) reported a significant direct response for a selected trait (live weight) and a correlated response for shell length in *O. chilensis*. In the same way, genetic correlations between size traits were high (close to unity) for *P. vannamei* (Perez-Rostro and Ibarra, 2003). Nevertheless, unfavourable genetic correlations between survival in the first 4 months and shell length until 2 years of age have been reported in the red abalone (Jonasson et al., 1999).

Another important point to take into consideration in a selection programme is the age at which the superior individuals should be selected as parents for the next generation. This relies on the phenotypic and genetic correlations between values of traits at near market size and values at earlier stages (larval or juvenile). Selecting at an early stage is of great advantage for the breeder since it reduces the cost of production. Nevertheless, some studies showed little or no

correlations between early growth and growth at later stages (Stromgren and Nielsen, 1989; Toro and Newkirk, 1990; Jonasson et al., 1999). In the case of low correlations of a trait between ages, the breeder needs to select the individuals as close to the market size as possible. Toro et al. (1995) reported that the selection of *O. chilensis* could be effective at 8 or 14 months. However, Collet et al. (1999) showed the influence of early growth on growth in later stages in *C. gigas* suggesting that selection for growth could be performed at an early stage.

II.5. Marker-assisted selection

Quantitative traits can be improved using the classical approach of selection experiments, based on the underlying quantitative genetics theory.

Another promising way to achieve this goal is to use marker-assisted selection. This approach relies on the development of linkage maps, i.e. genetic mapping of markers in the genome. The identification and isolation of markers flanking Quantitative Trait Loci (QTL) will then allow the selection of lines or individuals based on their genotypes for these markers. This is an indirect approach that consists of selecting for marker genotypes based on a close link between the marker and the QTL of interest. This search for QTL is of course justified for traits with a genetic basis, i.e. showing a significant heritability in classical quantitative genetics experiments. The marker-assisted selection method therefore requires the establishment of combined maps of markers and QTL. Ward et al. (2000) illustrate the potential improvement that marker-assisted selection can bring in *C. gigas*.

In the following sections, methods are given on the way to establish a linkage map of markers and the identification of QTL linked to markers. The implications are discussed in terms of improvement of quantitative traits through marker-assisted selection.

III- GENETIC MAPPING

Genetic maps show the ordering of marker loci on chromosomes and the relative distances between them. Their construction requires the use of highly polymorphic marker loci,

equally dispersed throughout the genome. Such maps represent a framework which allows the localisation of QTL and are the first step in the search for such loci.

III.1. Recombination frequencies. Mapping functions

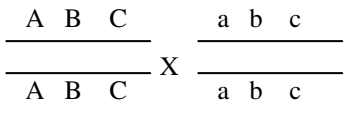
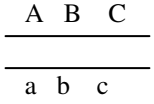
Chiasmata (singular, chiasma) represent areas along the chromosome where the paternal and maternal non-sister chromatids (belonging to homologous chromosomes) can be in tight association during the prophase of the first division of meiosis (meiosis I). They are associated with crossing-over or "recombination", i.e. reciprocal exchanges of portions of chromosomes around the breaking point. Crossing-over result in recombinant gametes, as opposed to parental gametes (chromatids non affected by the crossing-over). Each bivalent (a pair of homologous chromosomes during the first meiotic division) normally has at least one chiasma. The number of chiasmata occurring is roughly proportional to the length of the chromosome, but cytological analyses have revealed that any chromosome can have more than five or six chiasmata (Kearsey and Pooni, 1998).

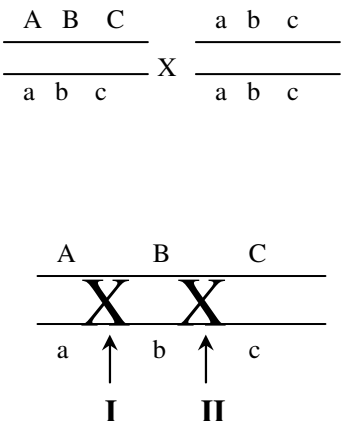
The frequency of recombinant gametes is a function of the distance separating two loci. It is obvious that the further apart on a chromosome two loci are, the greater the probability is that one or more cross-overs will occur between them; the observation of "recombinant" gametes will be highly frequent. On the contrary, when two loci are close together on a chromosome, it will be very unlikely that a chiasma occurs between them and hence many fewer recombinant gametes will be observed. This relationship between the distance separating two loci on a chromosome and the recombination frequency is the basis for the estimation of the genetic distance.

To illustrate how the ordering of markers and their relative distances can be estimated from the observed recombination frequency, consider three markers A, B and C (Figure 5). In this example the parental lines are homozygotes for different alleles at all the markers. The use of polymorphic markers is essential because otherwise the genotypes of the progeny cannot be unambiguously determined. The recombination frequency between two markers is estimated as the observed number of recombinant gametes divided by the observed total number of gametes. Hence, in the example given Figure 5, the recombination frequency between A and B is estimated

as 0.1925 ((14+7+24+32)/400), between A and C as 0.3025 and between B and C as 0.3900. The ordering of the three markers is then deduced: B, A and C.

The base unit of the genetic map is the centiMorgan (symbol cM). The length of chromosome which on average has one cross-over is defined as 50 cM long. A first approximation considers that the recombination frequency and the genetic distance in cM are equal. In our example, the distance between B and A would be 19.25 cM, between A and C 30.25 cM and between B and C 39 cM. However, this simple relationship is not valid over long intervals of chromosome. Indeed, as seen in the previous example, recombination frequencies are not additive because recombination frequency between B and C is not equal to the sum of the intermediate recombination frequencies. Recombination frequencies between two loci will plateau at 50% whereas genetic distances will increase linearly with the number of chiasmata (Figure 6).

<table>
<tr><td>Parental lines</td><td>  </td><td colspan="2">Homozygous for the 3 markers</td></tr>
<tr><td>F1</td><td>  </td><td colspan="2">Heterozygous for the 3 markers</td></tr>
<tr><td>Back-cross (Test-cross): cross between an F1 individual with the recessive homozygous line (abc).

Schematisation of the possible cross-overs occurring along the chromosome carrying the three markers A, B and C. Gametes generating from these F1 are either parental associations (p), or recombinant gametes (r). The recombinant gametes result either from a crossover in the area I, or in the area II or from a double crossover (I+II).</td><td>  </td><td>F1 gametes

ABC (p)
abc (p)
ABc (r II)
abC (r II)
Abc (r I)
aBC (r I)
AbC (r I+II)
aBc (r I+II)
Total</td><td>Number of progeny

103
120
51
49
14
7
24
32
400</td></tr>
</table>

Figure 5. Example of calculation of recombination frequencies in a backcross family, with three markers A, B and C (Kearsey and Pooni, 1996).

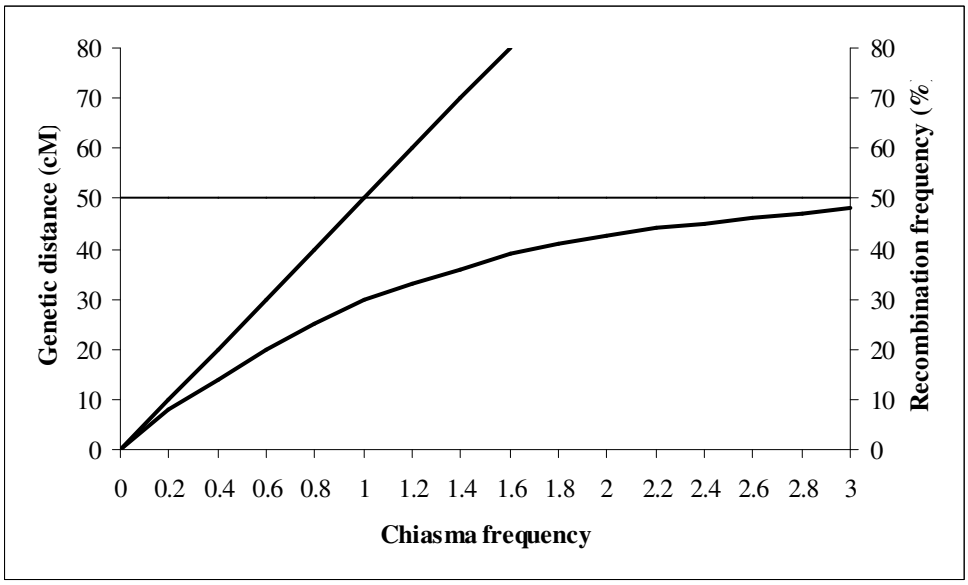


Figure 6. Genetic distance and recombination frequency as a function of chiasma frequency.

In contrast to genetic distances, recombination frequencies are not additive. The estimation of the genetic distance separating 2 markers is a function of the average number of cross-overs that occur between them. Therefore, various mapping functions have been developed to predict the genetic distance (in cM) from the observed recombination frequency.

Haldane's mapping function (Haldane, 1919) assumes no interference, i.e. the occurrence of a crossover in one region of the chromosome does not affect the frequency of crossovers in adjacent regions. Therefore, crossovers occur randomly and independently over the entire chromosome. The Haldane map distance (m_H) in cM can be estimated directly from the observed recombination frequency (c):

$$m_H = -\frac{\ln(1-2c)}{2} \times 100$$

For example, a recombination frequency of 0.16 will correspond to a genetic distance of 19.2 cM.

Another commonly used mapping function, developed by Kosambi (1944), takes into account the possible crossover interference in adjacent sites. The Kosambi map distance (m_K) is estimated as:

$$m_K = \frac{1}{4}\ln\left(\frac{1+2c}{1-2c}\right)\times 100$$

where c is the recombination frequency between 2 markers. In this case, a recombination frequency of 0.16 correspond to a genetic distance of 16.6 cM.

Both mapping functions give $m \cong c$ for small recombination frequency (<0.15), and for large m both approach c=0.5 (Lynch and Walsh, 1998).

III.2. Establishing framework maps

The example given in Figure 5 illustrated the use of a backcross design to determine the genetic distance between three markers and determine their relative order along a chromosome. This procedure can be repeated with other markers to map them with respect to the already mapped markers. However, the development of molecular markers based on DNA polymorphisms (such as microsatellite, AFLP or RFLP markers) leads to a high potential number of markers to be used. Therefore, the very large number of loci segregating simultaneously in a cross requires the use of computer software in order to group the markers into linkage groups and establish the overall map.

Several software programmes are presently available for genetic mapping. The most widely used are LINKAGE (Suiter et al., 1983), MAPMAKER (Lander et al., 1987), JOINMAP (Stam, 1993), MAPMANAGER QTX (Manly and Olson, 1999) and CRIMAP (Green et al., 1990).

MAPMAKER, MAPMANAGER, CRIMAP and JOINMAP software use the maximum likelihood approach to group markers into linkage groups and estimate the recombination frequencies between them. The grouping of markers into linkage groups relies then on the analysis of likelihood ratios or LOD scores (Logarithm of the Odds, see part IV.4. of this chapter). MAPMAKER has been primarily developed for the construction of primary genetic linkage maps from RFLP data. Two kinds of families can be used for this construction: either data from F2

intercrosses/backcrosses between homozygous inbred lines (experimental crosses) or data from two- or three-generation families in a natural population (analysis of grand-parents, parents and offspring). CRIMAP is very similar to MAPMAKER. The main advantage of CRIMAP comes from the coding of data. Indeed, both codominant and dominant markers can be handled in the same dataset. Contrary to MAPMAKER, JOINMAP has been designed to construct integrated genetic maps, using segregation information obtained from different laboratories, with different families, markers and experimental design. It helps to integrate linkage maps that have been obtained independently to construct a combined map. This software can be used for different types of crosses as well as for dominant and/or codominant markers; its applications seem wider. Moreover, the analysis is utterly automatic and so less subject to evaluation errors from the experimenter. MAPMANAGER QTX is a graphic, interactive program to map Mendelian loci as well as QTL using intercrosses, backcrosses or recombinant inbred strains in experimental plants or animals.

LINKAGE allows the analysis of both dominant and codominant markers. Contrary to software previously cited, LINKAGE uses contingency chi-square analyses to test linkage between all pairs of markers instead of maximum likelihood approach.

MAPCHART (Voorrips, 2002) is a software which allows the representation of genetic maps from linkage map data. It can be downloaded at http://www.joinmap.nl.

III.3. Genetic maps established for aquaculture species

The development of genetic linkage maps is particularly useful for the mapping of quantitative trait loci (QTLs) and for marker-assisted selection (Lander and Botstein, 1989).
Extensive genetic maps, based on a very high number of segregating markers, have been established for agricultural animals: sheep (de Gortari et al., 1998), chicken (Groenen et al., 2000), cattle (Kappes et al., 1997) and swine (Rohrer et al., 1996). Moreover, comprehensive genetic maps have been reported for human (Dib et al., 1996) and mouse (Dietrich et al., 1996) based on 7377 markers (microsatellites and RFLPs) and 5264 microsatellites respectively.

Because of the large number of genetic mapping studies, this literature review is restricted to aquaculturally important species to highlight the different kinds of markers and experimental designs that can be used to achieve the construction of a genetic map. Genetic linkage maps have been established for almost all major aquaculturally important species, including tilapia, catfish, salmon, rainbow trout, oysters and shrimps as well as for other fish species. The characteristics of these genetic maps and methods employed to construct them are summarised in Table 4 for shellfish species and Table 5 for fish species. It is striking that the studies concerning bivalve species are very scarce. A few studies reported the construction of a preliminary genetic map in *C. gigas* (Hubert and Hedgecock, 2004; Li and Guo, 2004). Apart from preliminary linkage relationship data between some allozyme loci (Beaumont, 1994a) no genetic mapping has been reported in mussels.

Various types of markers can be used for genetic mapping studies. Before the advent of molecular markers, genetic maps were built based on the segregation of allozyme loci (e.g. Thorgaard et al., 1983 in rainbow trout; Morizot et al., 1991 in *Xiphophorus spp*.). But these studies were limited by the amount of polymorphic loci available (8 and 76 respectively). Nevertheless, Morizot et al. (1991) established 17 multipoint linkage groups, with 56 (out of 76) allozyme loci mapped. May and Johnson (1989) established an allozyme composite linkage map composed of 18 LGs and 6 pseudolinkage groups, of several Salmonid species (charrs, trouts and salmons). The advance of molecular markers, based on DNA polymorphisms, allowed the construction of more complete genetic maps.

Genetic maps based on the segregation of microsatellite markers have been established for rainbow trout (Sakamoto et al., 2000), zebrafish (Gates et al., 1999; Shimoda et al., 1999; Kelly et al., 2000), tilapia (Kocher et al., 1998; Agresti et al., 2000), channel catfish (Waldbieser et al., 2001) and Japanese flounder (Coimbra et al., 2003). Moreover, microsatellite markers have been used to construct a genetic map of the cupped oyster *C. gigas* (Hubert and Hedgecock, 2004). Microsatellites are very polymorphic and codominant, features required in the context of genetic mapping. Nevertheless, saturated genetic maps require a very high number of microsatellites whose development is very time consuming and expensive.

AFLP (Amplified Fragment Length Polymorphism) markers are another kind of markers commonly used in genetic mapping. Their relatively easy and rapid development, without any prior knowledge of the genome of the species studied, explains their wide use. In shellfish, AFLPs represent the most commonly used markers and have been used to construct preliminary maps in various shrimp species (Moore et al., 1999; Wilson et al., 2002; Li et al., 2003), in the American oyster *C. virginica* (Yu and Guo, 2003) and in the cupped oyster *C. gigas* (Li and Guo, 2004). In fish species, maps have been established for tilapia (Kocher et al., 1998; Agresti et al., 2000), rainbow trout (Young et al., 1998), walking catfish (Poompuang and Na-Nakorn, 2004), Japanese flounder (Coimbra et al., 2003) and medaka (Naruse et al., 2000). But AFLPs also present disadvantages. First, they are dominant markers so the segregation of markers in a diploid family will be ambiguous. Then, there is a doubt on the transferability of AFLPs from one family to another, or one strain to another. Moore et al. (1999) reported that most AFLPs were consistent across families of *P. japonicus*, but not across families of *P. monodon* or *P. esculentus*. Moreover, the few common markers between the three segregating families of *P. monodon* lead to the construction of family-specific maps and a common map was then established based on the few common markers (Wilson et al., 2002). Some studies combine both types of markers. A few microsatellites can serve as anchor loci and allow the comparison of genetic maps obtained in different laboratories or between different families, strains or even species, while AFLPs markers, due to the high number of markers that can be obtained quite easily, allow the saturation of genetic maps (Kocher et al., 1998; Agresti et al., 2000; Coimbra et al., 2003).

RAPD (Randomly Amplified Polymorphic DNAs) markers have been successfully employed in the construction of genetic maps in zebrafish (Postlethwait et al., 1994; Johnson et al., 1996). Liu et al. (1999) tested the feasibility of RAPDs for genetic mapping in channel and blue catfish but poor reproducibility limited their application. In addition they concluded that the low levels of intraspecific variation found in the RAPD profiles exclude the use of RAPDs as an efficient system to generate molecular markers for gene mapping in an intraspecific mating plan.

ESTs (Expressed Sequence Tags) are sequences of nucleotides that have been derived from gene expression. These sequences are generated by sequencing of clones obtained from complementary cDNA libraries (issued by reverse transcription of messenger RNAs, mRNAs)

and represent the products of genes expressed in a particular tissue. The cDNAs libraries constructed are most of the time tissue-specific (Lehnert et al., 1999; Jenny et al., 2002). Options for mapping ESTs include intron-length polymorphisms, SSCPs (Single Strand Conformation Polymorphism) and in some cases genotyping of microsatellites present in these sequences. ESTs are an important type of markers in genetic mapping studies since they can act as anchor loci and then allow the construction of cross-species genetic maps with the ultimate comparison of maps between species (Lehnert et al., 1999). The ShrimpMap project is an international cooperation (http://www.aims.gov.au/pages/resaerch/shrimpmap/pages/sm-05.html) which aims to develop a framework genetic map of the giant tiger prawn *P. monodon*. In this context, identification of ESTs has begun. Lehnert et al. (1999) reported the isolation of ESTs generated from cephalothorax, eyestalk and pleopod libraries in *P. monodon*. An EST database, ShrimpEST, has been created. In the context of cross-species comparison of shrimp genome maps, such a database will be useful because Moore et al. (1999) found it difficult to transfer microsatellite loci from one penaeid species to another. A marine species database of ESTs is available on the Internet site http://www.marinegenomics.org. For example, 5586 sequences are available for *C. virginica* and 3971 sequences for *C. gigas*. Jenny et al. (2002) identified potential genes (ESTs) involved in a stress response in *C. virginica* by constructing a cDNA library issued from hemocytes. As ESTs result from the expression of genes, they are expected to be more conservative between species (compared to non coding DNA). These sequences are therefore very promising markers for the between-species comparison of genetic maps, with implications in terms of evolution of genomes (see Gates et al., 1999; Woods et al., 2000; Kelly et al., 2000 for the comparison between zebrafish, human and mouse genomes).

Inspection of Tables 4 and 5 reveals the use of different experimental designs while the classical mating schemes simply involve the analysis of either backcross, or F2 progeny. As a reminder, a backcross consists of the mating of a hybrid (heterozygous for all loci) with the homozygous recessive line while F2 progeny result from the cross of two F1 hybrids. Numerous studies exist on the analysis of inter-strain crosses. Some mapping panels consist of a three-generation pedigree (grand-parents, parents and offspring) of backcross families (Sakamoto et al., 2000; Naruse et al., 2000; Li and Guo, 2004), or F2 families (Shimoda et al., 1999; Li et al., 2003). Other mapping panels consist of two generation pedigree (parents and offspring)

(Coimbra et al., 2003). The use of inter-strain crosses increases the polymorphism of genetic markers and this increases the mapping efficiency because it increases the linkage disequilibrium in the segregating families. However, these backcross and F2 crosses involve the use of inbred lines, homozygous for all loci, whose development can be time consuming, depending of the life cycle duration of the species under study; therefore such crosses tend to be restricted to experimental populations. When studying natural populations, or when inbred lines are not available, individuals can be taken from the population, genotyped and mated in pairs to yield a number of full-sib families. In a particular family, any pair of segregating loci will represent either an F2 (if both parents are heterozygous for the pair of markers) or a backcross (if only one parent is heterozygous whereas the other is homozygous) (Kearsey and Pooni, 1998). This strategy has been used in several studies (Moore et al., 1999; Waldbieser et al., 2001; Hubert and Hedgecock, 2004; Wilson et al., 2002; Yu and Guo, 2003) and is the most commonly used experimental design in shellfish species.

Alternative strategies have been developed to increase mapping efficiency by increasing the number of polymorphic markers segregating in the families. For example, interspecies crosses have been achieved in tilapia (Agresti et al., 2000) and Liu et al. (1999) exploited the channel catfish x blue catfish hybrid system to construct a genetic linkage map. A further approach involves recombinant inbred lines (RI) that can be produced by multiple generations of sib mating or self fertilization. After sufficient generations of inbreeding, each RI will fix different associations of the progenitor's alleles (homozygosity achieved). However, this strategy is very time consuming and has been used mostly in plant species (e.g. Burr et al., 1988 in maize). Another very similar strategy, the production of doubled haploid (DH) individuals, can be produced from line hybrids in a single generation using androgenesis (Young et al., 1998 in rainbow trout) or gynogenesis (Kelly et al., 2000 in zebrafish). The production of DH individuals involves the activation of eggs by sterizoa (irradiated sperm), or activation of irradiated eggs by normal sperm, followed by suppression of meiosis II or first cleavage division to make DH gynogens (all female chromosomes) or androgens (all male chromosomes). Using double haploids simplifies linkage analysis and allows the use of dominant markers since the genotype of the offspring can be unambiguously determined. Finally, another experimental design used in fish species is a haploid mapping panel, which consists of the analysis of a single parent (usually

female) and its haploid offspring (derived by gynogenesis). Genetic maps based on the segregation of markers in a haploid family, have been established in walking catfish (Poompuang and Na-Nakorn, 2004), zebrafish (Postlethwait et al., 1994; Johnson et al., 1996; Gates et al., 1999), Atlantic salmon (Slettan et al., 1997) and tilapia (Kocher et al., 1998). As when using double haploids, it is again easier to analyse the segregation of dominant markers in a haploid family because identification of homozygotes and heterozygous is possible. However, using DH individuals present a considerable advantage compared to the use of haploid embryos since doubled embryos can survive until adult stages. Because of the higher amount of DNA that can be extracted, it becomes possible to map numerous markers in a single cross, diminishing the need to integrate maps produced in different mapping panels (Kelly et al., 2000).

There are some general features of genetic maps that should be noted. Firstly, clustering of markers has been highlighted in different studies. The clustering of markers has been more frequently found with AFLPs markers; for example in tilapia (Kocher et al., 1998; Agresti et al., 2000) or in rainbow trout (Young et al., 1998). According to these studies, AFLP clusters identify the heterochromatic regions associated with centromeres and indicate the lack of recombination in that region of chromosomes. In the same way, clustering of microsatellite markers has been reported in zebrafish genome (Shimoda et al., 1999). These grouping of markers in some centromeric regions were once again related to the suppression of recombination near centromeres (Johnson et al., 1996).

Secondly, chiasma interference is common in fish species. This phenomenon has been reported for zebrafish (Johnson et al., 1996), medaka, channel catfish, rainbow trout (Thorgaard et al., 1983) and tilapia (Kocher et al., 1998). For this reason Kosambi's mapping function was used in most studies of genetic mapping in fish species.

Thirdly, large sex-specific differences in recombination rates have been reported in several studies. Higher recombination rates in females have been found in rainbow trout (Sakamoto et al., 2000; 3.25 female: 1 male), channel catfish (Waldbieser et al., 2001; 3.18 female: 1 male), zebrafish (Knapik et al., 1998), *C. virginica* (Yu and Guo, 2003), *P. monodon* (Wilson et al., 2002). These female: male ratios in recombination rates were obtained by pairwise

comparisons of average spacing between markers common to male and female parents. In these studies, average marker spacing is slightly less for the male map than for the female map. The linear relationship between linkage distance and recombination rate implies that the recombination rate is slightly lower in male compared to female. Similar recombination rates were reported between male and female in *P. japonicus* (Li et al., 2003) since average spacing of markers is similar between males and females. At the other extreme, a study on the Japanese flounder (Coimbra et al., 2003) reported much higher recombination rates in males (7.4 male: 1 female). Usually, genetic length of maps should reflect the differences in recombination frequency between sexes, i.e. higher recombination rates in females should be associated with longer female genetic map length. Such associations have been reported in humans (Dib et al., 1996), mouse (Dietrich et al., 1996), rainbow trout (Sakamoto et al., 2000), *C. virginica* (Yu and Guo, 2003) with bigger female map length linked to higher recombination rate in female. However, some studies do not report such congruence. For example, Coimbra et al. (2003) reported similar total genetic map length between male and female despite higher recombination rate in male. In this case, the female was heterozygous for a higher number of markers, meaning that more markers were assigned on the map, resulting in an expansion of the female map. The male genetic map in *P. japonicus* is far longer (1781cM) than the female one (1026cM), despite similar average spacing (Li et al., 2003) but this can again be explained by the presence of more mapped markers in the male parent, resulting in a higher apparent genome coverage. Another example of sex-discrepancy between genetic map lengths due to the number of informative markers is the study of Agresti et al. (2000) in tilapia: the female genetic map covers 514cM with 62 markers mapped, whereas the male map covers 1632cM with 214 mapped markers. Therefore, the number of polymorphic markers influences the length of the genetic map. Because the experimental design used can determine the number of polymorphic markers, experimental design indirectly influences genetic map length. In one of the examples cited above (Agresti et al., 2000), the male parent was an F1 hybrid issued from a cross between two species, resulting in more polymorphic markers and hence a longer male map length. Efforts should be made to compare male and female maps obtained in a similar manner, e.g. with the same number of markers. Despite the direction of sex-specific differences in recombination frequency, the cause of such discrepancy is unknown. In Salmonids species only, the repression of male

recombination is due to the occurrence of multivalent pairings (due to the ancient tetraploidy of the Salmonids) in male meiosis only: the formation of such pairings seems to inhibit crossovers.

Fourthly, depending on the studies, a discrepancy of genetic length can often be noticed. A striking example is the zebrafish genetic maps where adding markers to a map increases its length. Postlethwait et al. (1994) reported a length of 2317cM with 424 markers, Johnson et al. (1996) 2790cM with 652 markers and Kelly et al. (2000) 3011cM with 1364 markers. This is logical because adding markers to the map is bound to increase the genome coverage. Total genetic length (extrapolated by taking into account average spacing between markers, gaps and telomeres) obtained in different studies for the same species are more similar, e.g. 2720-3000cM for the zebrafish. However, Yu and Guo (2003) hypothesized that low marker density is the primary cause for the longer than expected genetic length observed in *C. virginica*.

III.4. Conclusion

The construction of genetic maps relies on the analysis of segregation of markers among progeny. Various kinds of markers can be used in genetic mapping studies: AFLPs, RFLPs, microsatellites, ESTs. Ideally, several types of markers can be combined. AFLP development is rapid and requires no knowledge of the genome, so AFLPs can be used to saturate genetic maps. However, they are not always reliably transferable from one family or laboratory to another. The use of a few microsatellites (common between families and sometimes species) can serve as anchor loci and allow the construction of integrated genetic map, using information derived from several families or mapping panels. Moreover, the use of ESTs, which are conserved across species, allow the construction of cross-species maps and the comparison of the evolution of genomes.

Different experimental designs have been developed, the most common being backcross and F2 families in experimental populations and full-sibs families in natural populations. Other strategies have been used in fish species to increase mapping efficiency: double haploid, haploid mapping panel, recombinant inbred lines. Because of the technical complexity required and biological constraints of the production of haploids as well as the high load of work for the development of RI lines, the latter strategies are not readily applicable to shellfish species in the

near future. Nevertheless, artificial gynogenesis has been achieved in the cupped oyster *C. gigas* (Guo et al., 1993; Guo and Gaffney, 1993) so the production of haploids could be an efficient way for achieving genetic mapping in shellfish. In these species, efficiency in genetic mapping should be improved by the development of inbred lines, by inter-strain crosses or by careful selection of progenitors used in the crosses (based on their genotype to maximize linkage disequilibrium between the parents of the cross).

Genetic maps, based on the segregation of a high number of polymorphic markers, are constructed using specific software such as MAPMAKER, JOINMAP, MAPMANAGER and CRIMAP. They are based on maximum likelihood procedure and the computation of LOD score to determine the maximum likelihood order and genetic distances between pairs of markers.

Once genetic maps have been constructed, they provide a framework for further studies, particularly the identification and location of QTLs.

Table 4. Review of primary genetic maps obtained for various shellfish species. n: haploid number of chromosomes; LGs: linkage groups; θ: recombination frequency LOD score (Logarithm of the Odds); COMPARE, ORDER, RIPPLE, ERROR DETECTION, TRY: commands of MAPMAKER software (see Chapter 3 for details)

Species	Genetic map features	Markers used	Experimental design	Software used	References
Penaeus japonicus Kuruma prawn (n=43)	44 LGs (14 with more than 3 markers) 129 markers mapped Genetic length 1276 cM Average spacing 15 cM Genome coverage 57%	246 AFLP loci (15 primer sets)	One full-sib family of 41 offspring DNA available from progeny, parents and grandparents (three-generation pedigree).	**MAPMAKER**, F2 intercross. Initial grouping based on LOD score 3.5 and maximum recombination frequency of 0.25. Then, multipoint analysis.	Moore et al. (1999)
	Male linkage map: 43 LGs (27 major) 227 markers mapped Genetic length 1781 cM Average spacing 9.7 cM Female linkage map: 31 LGs (16 major) 125 markers mapped Genetic length 1026 cM Average spacing 10.9 cM	401 AFLP loci (54 primer sets): 251 ♂, 150 ♀.	One F2 cross family (HLxLH). Grandparents HH and LL. (LL and HH: low and high line for size criterion) Initial step: 46 progeny (top/bottom 6% of size distribution). Second stage: 102 progeny (top/bottom 8% of size distribution).	**MAPMAKER**, F2 backcross model Initial grouping based on LOD score 5.0 and maximum recombination frequency of 0.25. Then, multipoint analysis. Less-stringent criteria (LOD=3.5; θ=0.30). Haldane's mapping function.	Li et al. (2003)
Penaeus monodon Black tiger shrimp (n=44)	Common map (3 families) 19 LGs (8 major) 63 markers mapped Genetic length 1412 cM Average spacing 22 cM Male, female maps for each family done too	673 AFLP loci (23 primer sets) 116 AFLP loci common to the three families	Three-generation pedigreed families Three full-sibs families (for each, 4 grandparents, 2 parents and the offspring are analysed).	**MAPMAKER**, F2 backcross model Initial grouping based on LOD score 5.0 and maximum recombination frequency of 0.25. Then, multipoint analysis. Less-stringent criteria (LOD=3.5; θ=0.30). Haldane's mapping function.	Wilson et al. (2002) ShrimpMap project
Penaeus chinensis Chinese shrimp (n=44)	Male linkage map: 36 LGs 194 markers mapped Genetic length 1737.3 cM Average spacing 11.0 cM Genome coverage 68.1% Female linkage map: 35 LGs 197 markers mapped Genetic length 2191.1 cM Average spacing 13.5 cM Genome coverage 69.6%	501 AFLP loci (88 primer sets) 247 ♂, 254 ♀.	One F1 family issued from cross between 2 Chinese shrimps with different genetic bases (♀ from a selected population, ♂ from the wild). Pseudo-test cross strategy.	**MAPMAKER**, F2 backcross model Initial grouping based on LOD score 40 and maximum genetic distance of 45 cM. COMPARE command for small groups (≤8). ORDER command for larger groups. ERROR DETECTION command used. Kosambi's mapping function.	Li et al. (2006b)

Penaeus vannamei White shrimp (n=44)	Female linkage map 51 LGs 212 markers mapped Genetic length 2771 cM Average spacing 17.1 cM Genome coverage 62% Male linkage map 47 LGs 182 markers mapped Genetic length 2116 cM Average spacing 15.6 cM Genome coverage 59%	741 AFLP loci (103 primer sets)	A family derived from a commercial stock under domestication for five generations. Only genotype of the female known. 42 progenies analysed. Pseudotestcross strategy used.	**MAP MANAGER QTX.** Initial grouping based on LOD score 3.0. Then, unlinked markers were assigned to the framework linkage groups with a LOD score of 2.0. Kosambi's mapping function.	Perez et al. (2004)
Haliotis discus hannai Ino Pacific Abalone (2n=36)	Male linkage map 19 LGs 94 markers mapped Genetic length 1365.9 cM Average spacing 18.2 cM Genome coverage 73.1% Female linkage map 22 LGs 119 markers mapped Genetic length 1773.6c M Average spacing 18.3 cM Genome coverage 74.9%	135 AFLP loci 5 RAPD loci 6 microsatellites 1 sex marker 230 AFLP loci 5 RAPD loci 6 microsatellites 1 sex marker	F1 mapping family issued from the single-pair mating of two abalones from 2 geographically distant populations (female from Japanese population and male from Chinese population). The family studied consists of the two parents and 86 progeny.	**MAPMAKER**, F2 backcross model Initial grouping based on LOD score 4.0 and maximum genetic distance of 35 cM for determining linkage groups. Then, three point and multipoint analysis. Final order of each group tested with RIPPLE function. ERROR DETECTION command used. Kosambi's mapping function.	Liu et al. (2006)
Strongylocentrotus Sea urchin (n=21)	Male linkage map 23 LGs 199 markers mapped Genetic length 2614.8 cM Average spacing 15.4 cM Genome coverage 79.1% Female linkage map 24 LGs 194 markers mapped Genetic length 2988.3 cM Average spacing 17.1 cM Genome coverage 78.1%	897 AFLP loci (42 primer sets) 446 ♂, 451 ♀	Interspecific cross between *S. nudus* (♀) and *S. intermedius* (♂). 60 F1 progeny (6 month-old). 2 way pseudo-test cross strategy.	**MAPMAKER**, F2 backcross model Initial grouping based on LOD score 3.0 and maximum recombination frequency of 0.30. COMPARE command for small groups (≤8). ORDER command for larger groups. Then, lowering LOD score criteria to 2.0 to map additional markers on the framework map. Final order of each group tested with RIPPLE function. ERROR DETECTION command used. Kosambi's mapping function.	Zhou et al. (2006)

Crasssostrea virginica Eastern oyster (n=10)	Male linkage map 12 LGs 114 markers mapped Genetic length 647.4 cM Average spacing 6.3 cM Genome coverage 83.7%	153 AFLP loci 3 microsatellites 2 EST markers	Reference family produced by single-pair mating of the Rutgers NEH strain (selected for resistance against two diseases, MSX and Dermo). The family studied consists of the two parents and 81 progeny.	**MAPMAKER** Initial grouping based on LOD score 3.0 and maximum distance of 38 cM. Small groups ordered with multipoint analysis, large groups with the three-point analysis. Then lowering of LOD score to ≥2.0 to add markers. Kosambi's mapping function.	Yu and Guo (2003)
	Female linkage map 12 LGs 84 markers mapped Genetic length 904.3 cM Average spacing 12.6 cM Genome coverage 84.2%	129 AFLP loci 3 microsatellites 1 EST marker			
Crassostrea gigas Pacific oyster (n=10)	Male linkage map 10 LGs 88 markers mapped Genetic length 616.1 cM Average spacing 8.0 cM Genome coverage 79%	115 microsatellites	Three mapping families (double-hybrid crosses among 4 filial lines)	**MAPMAKER** Initial grouping based on LOD score 3.0. Linkage assessed separately for the female and male parents of each mapping family. Error detection command used. Kosambi's mapping function. Consensus map for each sex established by Compiling data from the three families.	Hubert and Hedgecock (2004)
	Female linkage map 10 LGs 86 markers mapped Genetic length 770.5 cM Average spacing 10.4 cM Genome coverage 70%				
	Male linkage map 10 LGs 96 markers mapped Genetic length 758.4 cM Average spacing 8.8 cM Genome coverage 81.3%	349 AFLPs (191 ♀, 158 ♂) 44 primer sets screened, 17 selected for mapping	Reference family with 73 progeny Reference family: backcross of an interstrain hybrid to one of the parental strain	**MAPMAKER** Initial grouping based on LOD score 3.0 and maximum distance of 30 cM. Once LGs determined, ordering of markers by a three-point analysis. Error detection command used. Kosambi's mapping function.	Li and Guo (2004)
	Female linkage map 11 LGs 119 markers mapped Genetic length 1030.7 cM Average spacing 9.5 cM Genome coverage 82%				

	Male linkage map 19 LGs 94 markers mapped Genetic length 1511.4cM Average spacing 20.1cM Genome coverage 66.6% Female linkage map 20 LGs 97 markers mapped Genetic length 1610.2cM Average spacing 20.9cM Genome coverage 66%	783 AFLP loci (21 primer sets)	Mapping family consists of 51 F1 progeny derived from a biparental cross (wild population). Markers selected were in 3:1 or 1:1 ratio.	**MAPMAKER/EXP** F2 backcross model Initial grouping based on LOD score 3.0 and maximum distance of 30cM. Then multipoint analysis. Once framework LGs established, use of less stringent criteria (LOD of 2.5). Error detection command used. Kosambi's mapping function.	Wang et al. (2004)
Chlamys farreri Zhikong scallop (n=19)	Male linkage map 20 LGs 197 markers mapped Genetic length 1630.7 cM Average spacing 9.2 cM Genome coverage 81.7% Female linkage map 19 LGs 166 markers mapped Genetic length 1503.9cM Average spacing 10.2cM Genome coverage 79.6%	667 AFLP loci (217 ♀, 270 ♂, 213 segregating in both parents) 32 primer sets.	Mapping family derived from a cross between an individual from a Japanese population and another from a Chinese population. Family chosen according to polymorphism level. The family studied consists of the two parents and 93 progeny.	**MAPMAKER.** Two-way pseudo-testcross strategy. Two data sets obtained for the maternal and paternal parents. Initial grouping based on LOD score 3.0 and maximum distance of 30cM to map 1:1 segregating markers. Then, multipoint analysis. Finally addition to the framework map of unlinked and distorded markers (LOD of 2.7 and maximum distance of 35cM). Kosambi's mapping function.	Li et al. (2005)
	Male linkage map 23 LGs 166 markers mapped Genetic length 2468 cM Average spacing 14.9 cM Female linkage map 25 LGs 198 markers mapped Genetic length 3130 cM Average spacing 14.6 cM Consensus map 5 LGs Genetic length 431.2 cM Average spacing 17.2 cM	603 AFLP loci (76 primer sets) 329 ♀, 274 ♂, 100 segregating in both parents	Inter-population hybrid F1 family (♀ from wild Chinese population x ♂ from wild Japanese population). 60 F1 progeny (one year-old).	**MAPMAKER/EXP** F2 backcross model Initial grouping based on LOD score 3.0 and maximum distance of 50cM (sex-specific). COMPARE command for small groups (≤7). COMPARE, TRY commands for larger groups. Then markers heterozygous in both parents were analysed with an F2 inter-cross model, for establishing a consensus map. Lod score was set to 3.0 and then 2.0. RIPPLE command. ERROR DETECTION command used. Kosambi's mapping function.	Wang et al. (2005)

Table 5. Review of primary genetic maps in fish aquaculture species. See Table 4 for main abbreviations. STSs: Sequence Tagged Sites; IRSs: Internal Repeat Sequences.

Species	Genetic map features	Markers used	Experimental design	Software used	References
Oncorhynchus mykiss Rainbow trout (n=29-32)	42 LGs (31 major) 465 markers mapped Genetic length 1997.5cM Genome coverage 76% Clustering of AFLPs	332 AFLP loci 96 VNTR loci 40 SINEs loci 5 RAPD loci 2 microsatellites Sex	76 doubled haploid individuals produced by androgenesis were used in the segregation analysis. (Parent: hybrid resulted from the cross of two homozygous parental lines produced by androgenesis).	**MAPMAKER**, F2 intercross. Initial grouping based on LOD score 3.0 and maximum recombination frequency of 0.35. Then, analysis done using a minimum LOD Score of 4.0 and θmax=0.30 to detect anomalous linkages. Error detection function. Kosambi's mapping function.	Young et al. (1998)
	Male and female maps 29 LGs ♀ map length 1000 cM ♂ map length 1350 cM	191 microsatellites 7 allozyme loci 3 RAPD loci 7 ESTs	Three backcross families (F1 hybrid males x pure strain females).	**LINKMFEX** (Danzman, unpublished). Map distances assigned assuming complete interference. Centromere-linkage map: centromere-markers identified for 15 LGs with half-tetrad analysis.	Sakamoto et al. (2000) SalMap project
	30 major LGs 1314 markers mapped Genetic length 4359 cM Average spacing 7.4 cM	973 AFLP loci 226 microsatellites 72 VNTR 38 SINE 29 known genes 12 minisatellites 5 RAPD loci 4 allozymes	Double haploids produced by androgenesis (Male parent: F1 hybrid resulted from cross between two isogenic lines).	**MAP MANAGER QT** Grouping based on LOD score 10.and minimum distance of 0 cM. Remove of the representative marker for each LG for subsequent analysis. **MAPMAKER** Initial grouping based on LOD score 3.0 and maximum recombination frequency of 0.35. Then increasing the minimum LOD (4 and 5). Kosambi's mapping function.	Nichols et al. (2003)
	Male genetic map 28 LGs 257 markers mapped Genetic length 1019 cM Average spacing 3.9 cM	185 AFLP loci 72 microsatellites	Inter-strain cross. Cross between steelhead and rainbow trout: obtention of several F1 families. Then backcross families (F1 x rainbow trout). For the linkage analysis, only one family was chosen (binomial distribution of survivorship).	**MAPMAKER** Grouping based on LOD score 3.0 and a maximum distance of 50cM. AFLPs and microsatellites scored as dominant markers.	Rodriguez et al. (2004)
	Female genetic map 45 LGs 236 markers mapped Genetic length 2041 cM Average spacing 8.6 cM	164 AFLP loci 72 microsatellites			
Cyprinus carpio L. Common carp (n=50)	50 LGs 268 markers mapped Genetic length 4111 cM	110 microsatellites 105 genes 57 RAPD	Haploid gynogenesis. Analysis of a single haploid family: the grand-parents, the hybrid female (derived from cross between a female common carp with a male Boshi carp) and 80 haploid embryos.	**MAP MANAGER.** Genotype data entered in two phases.	Sun and Liang (2004)

Salmo salar Atlantic salmon	Male genetic map 31 LGs 251 markers mapped Genetic length 103 cM Female genetic map 33 LGs 230 markers mapped Genetic length 901 cM	473 AFLP loci (82 primer sets) 54 microsatellites	One paternal half sib family (two full sib families, with 69 and 67 offspring). Norwegian breeding population.	First two-point linkage analysis performed for each pair of informative markers: obtaining LOD scores and recombination fractions for male and female. Then, **JOINMAP 3.0.** Male data set: grouping of markers into LGs based on LOD score 4.0. Corresponding LGs identified in the female data set. Finally, lowering LOD score to 3.0 for adding markers to the framework. Kosambi's mapping function.	Moen et al. (2004b)
Ictalurus punctatus Channel catfish (n=29)	32 LGs 262 markers mapped Genetic length 1958 cM Average spacing 8.7 cM Genome coverage 87%	293 microsatellites (11 type I and 282 type II)	Analysis of two reference families: 2 parents and 72 full sib offspring per family.	**CRI-MAP** Two-point linkage analysis (LOD score of 3.0). Then multipoint analysis to determine the best order of markers. Comparison of recombination rates between males and females fro paired loci.	Waldbieser et al. (2001)
	Sex-average map 44 LGs (27 major) 418 markers mapped Genetic length 1593 cM	607 AFLP loci (64 primer sets)	Interspecies cross. F1 hybrid (channel x blue catfish) backcross either to channel or blue catfish. Reciprocal backcross made (F1 used as sire or dam). 8 backcross families produced; 71 offspring genotyped in each family.	**MAPMAKER** Initial grouping based on LOD score 3.0 and maximum recombination frequency of 0.30. First defining the most probable marker order of the most informative subset of each LG, then assigning additional markers.	Liu et al. (2003)
Clarias macro cephalus Walking catfish (n=27)	31 LGs 146 markers mapped Genetic length 2037 cM Average spacing 15.2cM	195 AFLP loci (47 primer sets)	Haploid gynogenesis. Analysis of a single haploid family: the female and 79 haploid embryos.	**MAPMAKER**, F2 backcross Initial grouping based on LOD score 3.0 and 50cM as maximum distance between linked markers. Kosambi's mapping function.	Poompuang and Na-Nakorn (2004)
Oreochromis Niloticus Tilapia (n=22)	30 LGs 162 markers mapped Genetic length 704 cM Average spacing 4.3 cM Genome coverage 70%	62 microsatellites 112 AFLP loci	Haploid gynogenesis : a single female was fertilized by irradiated sperm. Haploid embryos were collected at 2-3 days. Family studied consists of the female and 41 haploid embryos.	**MAPMAKER** Initial grouping based on LOD score 3.0. Kosambi's mapping function.	Kocher et al. (1998)
	Female genetic map 14 LGs 62 markers mapped Genetic length 514 cM	17 microsatellites 61 AFLP loci	Interspecies cross. Analysis of 63 progeny issued from a single mating between a female *Om* and an F1 hybrid male (*Oa*xR*On*). Three species are involved in the cross: *O. mossambicus* (*Om*), *O. aureus* (*Oa*) and Red *O. niloticus* (R*On*).	**MAPMAKER** Grouping based on LOD score 3.0. Kosambi's mapping function. **MAP MANAGER XP** Visual inspection of recombination events: detection of genotyping errors.	Agresti et al. (2000)
	Male genetic map 24 LGs 214 markers mapped Genetic length 1632 cM	62 microsatellites 167 AFLP loci			

Paralichthys olivaceus Japanese Flounder (n=23)	Male genetic map 25 LGs 223 markers mapped Genetic length 741.1 cM Average spacing 8 cM Genome coverage 64% Female genetic map 27 LGs 294 markers mapped Genetic length 670.4 cM Average spacing 6.6 cM Genome coverage 57%	82 microsatellites 149 AFLP loci 101 microsatellites 203 AFLP loci	Inter-strain cross. Parents: male produced by gynogenesis (strain A), female (strain B). Analysis of 44 F1 hybrids from this cross.	**MAP MANAGER QT** All loci scored as dominant markers. Grouping based on LOD score 3.0.	Coimbra et al. (2003)
Oryzias latipes Medaka (n=24)	24 LGs 633 markers mapped Genetic length 1354.5cM	488 AFLP loci 76 ESTs 28 RAPDs 34 IRSs 4 STSs 4 phenotypic	Inter-strain cross. Analysis of 39 progeny issued from the backcross: female AA2 crossed with the hybrid male (AA2xHNI).	**MAPMAKER** Grouping based on LOD score 3.5 and maximum recombination frequency of 0.35.	Naruse et al. (2000)
Seriola sp. Yellowtail (n=24)	Female genetic map 25 LGs 175 markers mapped Genetic length 473.3 cM Average spacing 2.7 cM Genome coverage 52% Male genetic map 21 LGs 122 markers mapped Genetic length 584.3 cM Average spacing 4.8 cM Genome coverage 34%	217 microsatellites (200 informative; 180 for female, 139 for male and 97 common to both parents)	Interspecies cross. Analysis of 90 progeny issued from a single mating between a female *S. quinqueradiata* and a male *S. lalandi*. Mapping panel consists of two generation-pedigrees, the two parents and the 90 F1 progeny.	**MAP MANAGER QT** Grouping based on LOD score 3.0. Haldane's mapping function.	Ohara et al. (2005)
Danio rerio Zebrafish (n=25)	29 LGs 424 markers mapped Genetic length 2317 cM Average spacing 5.8 cM Genome coverage 85%	405 RAPDs 16 microsatellites	Haploid gynogenesis. Analysis of a single hybrid female (DAR/AB) and 94 haploid offspring. Mapping of mutations using bulked segregant analysis.	**MAPMAKER** Grouping based on LOD score 3.0. Kosambi's mapping function.	Postlethwait et al. (1994)

Danio rerio Zebrafish (n=25)	Composite map 25 LGs 652 markers mapped Genetic length 2790 cM Average spacing 4.3 cM Genome coverage 96%	Markers of Postlethwaith *et al.* (1994) Add of 235 new RAPDs	1- Centromere linkage analysis: identification of markers closely linked to the centromeres by analysis of half-tetrads (females used: C32xDAR and C32xSJD) 2- Consolidation of existing genetic map by adding 235 markers. Haploid mapping panel composed of an hybrid female (C32xSJD) and 96 haploid embryos.	**MAPMAKER** Grouping based on LOD score 3.0. Kosambi's mapping function.	Johnson et al. (1996)
	Sex-averaged map 25 LGs 705 markers mapped Genetic length: 2100 cM Average spacing 5 cM Genome coverage 89.4%	705 microsatellites	Inter-strain cross. Mapping panel consists of three-generations pedigree: grandparents (AB female, IN male), parents (F1s) and 44 F2 progeny	**MAPMAKER** Initial grouping based on LOD score 5.0. Then, analysis repeated with less stringent criterion of LOD score 3.5. Error detection function. Confirmation of the map by centromere analysis (half-tetrad analysis on somatic-cell hybrid panel)	Knapik et al. (1998)
	Sex-averaged map 25 LGs 2000 markers mapped Genetic length 2145 cM Average spacing 1.2 cM Genome coverage 93%	Microsatellites	Inter-strain cross. Mapping panel consists of three-generations pedigree: grandparents (AB female, IN male), parents (F1s) and 44 F2 progeny. Centromere identified by half-tetrad analysis.	**MAPMAKER** Grouping based on LOD score 5.0. Then, LOD score of 4.0. Error detection function.	Shimoda et al. (1999)
	25 LGs 389 markers mapped Genetic length 2894 cM Genome coverage 99% Mapping of genes and ESTs combined with the microsatellite map of Knapik *et al.* (1998)	115 new SSCPs polymorphisms (cDNAs/ESTs, STSs) 275 previously mapped markers (microsatellites, genes, STSs)	Haploid mapping panel: analysis of a single hybrid female (TüxTL) and 48 haploid progeny.	**MAP MANAGER** First grouping of markers by maximising the LOD score and minimising the number of double crossovers. **MAPMAKER** Confirmation of the map order. Creation of map graphics and estimation of map distances using Kosambi's mapping function.	Gates et al. (1999)
	25 LGs 1364 markers mapped Genetic length 3011 cM	642 new SSCPs polymorphisms (genes and ESTs) 593 microsatellites (already mapped) 129 genes and ESTs (already mapped)	Homozygous diploid mapping panel (double haploids produced by gynogenesis then heat shock to restore diploidy). Two F1 hybrid females (C32xSJD) and 42 F2 double haploids (19 and 23 individuals from each female). C32 and SJD are two inbred strains.	**MAP MANAGER** First grouping of markers by maximising the LOD score and minimising the number of double crossovers. **MAPMAKER** Confirmation of the map order. Creation of map graphics and estimation of map distances using Kosambi's mapping function.	Kelly et al. (2000)

IV- QTL ANALYSIS AND MAPPING

IV.1. QTL: definition and features

Quantitative Trait Locus (QTL) refers to an individual locus that explains a certain proportion of the phenotypic variance of the trait. The plural will be noted "QTLs". For a better understanding of genetic determinism of a quantitative trait, it is therefore of primary importance to identify the number of genes (QTLs) contributing to quantitative variation. Several approaches have been developed to achieve this goal: i) detection of major genes relying on multimodal distribution and complex segregation analysis and ii) QTL mapping.

However, this definition is too restrictive. In fact, a QTL represents more a segment of chromosome affecting the trait rather than a single locus (Falconer and Mackay, 1996). This is due to the limit of resolution of QTL mapping. Indeed, QTLs closely linked to each other may appear as only one if in association. The classical limit of resolution of genetic mapping corresponds to a map distance of approximately 20cM. A QTL will then be a segment of chromosome of this length, and may contain one or more loci affecting the trait.

It has been shown in several studies (reviewed by Falconer and Mackay, 1996) that QTL loci segregate following Mendelian rules, exhibit additive effects of various magnitude and show various degree of dominance (from additivity to complete dominance or overdominance). Moreover, interactions between QTLs (epistasis) have been shown, as well as correlation between QTLs and QTL by environment interaction.

IV.2. Estimation of number of major genes influencing a trait

The first studies were based on phenotypic data to infer presence of major genes. Indeed, in some cases, most of the variation of the quantitative trait could be explained by the segregation of a few major genes. Several methods of detection of major genes were developed. The most simple rely on features of the phenotypic distribution: multimodal distribution or departure from normality could be taken as an indicator of the presence of a major gene. Another approach was based on offspring/parent resemblance. Under polygenic inheritance (no major gene), the mean

of offspring will resemble more closely the mid-parent value than single parents. On the contrary, when a major gene is segregating, the mean of offspring will resemble more closely to one of the parents. Carmelli et al. (1979) proposed a class of indices, major-gene indices (MGI), to reveal the presence of major genes. The most powerful approach, complex segregation analysis, relies on quantitative genetics. This approach consists of finding maximum likelihood estimates (MLEs) of parameters for a series of increasingly complicated hypotheses: a pure environmental model, a single gene model, a polygene model and the full mixed model. The significance of each hypothesis is tested by comparing the likelihood of the data, given maximum likelihood estimates of model parameters, with that calculated assuming the appropriate null hypothesis for which the tested parameters are set to zero, using a likelihood ratio test. Other hypotheses can be tested, such as common-family effects or a polygenic background in addition to the major gene. Every time, testing these hypotheses requires development of suitable models which account for these effects. The likelihood functions are detailed in Lynch and Walsh (1998).

IV.3. Principle of QTL mapping

Sax (1923) was the first to establish the basis of QTL mapping, relying on the associations between Mendelian markers and quantitative traits in crosses between inbred lines. Indeed, he demonstrated linkage between a monogenic trait (colour pigment of seed) and a quantitative trait (weight) in crosses between two inbred lines of beans *Phaseolus vulgaris*. This study was then the first to show linkage of genes controlling quantitative variation with single gene markers. Thoday (1961) developed this basic theme by using single gene markers scattered throughout the genome to map individual polygenes controlling quantitative variation and to detect their effects.

The principle of QTL mapping was thus laid out and identification of QTLs is based on linkage disequilibrium between alleles at a marker locus and alleles at the linked QTL. This notion of disequilibrium is of fundamental importance. Indeed, disequilibrium creates marker-trait associations, i.e. different marker genotypes having different expected values for characters influenced by QTLs linked to these markers. Therefore, QTLs mapping requires to analyse segregation of markers in populations showing disequilibrium: either crosses between inbred lines (fixed for different alleles) or analysis of relatives. Mapping QTLs consists of genotyping individuals at the marker locus and scoring their phenotype for the quantitative trait. The

presence of a QTL linked to the marker is inferred if there is a difference in phenotypic value among marker genotype classes. Marker loci can be considered singly or simultaneously.

Mapping QTLs requires i) the establishment of a linkage map of polymorphic markers located throughout the genome and ii) variation for the quantitative trait within or between populations. With the advent of molecular biology tools, molecular markers (highly polymorphic, abundant, neutral and codominant) have supplanted phenotypic markers previously used (Sax, 1923 used seed pigmentation as a marker). The use of these phenotypic markers was not very powerful as phenotype for these gene markers could hide and interfere with the quantitative trait. Therefore, the use of molecular markers was a fundamental step towards successful mapping of QTLs.

Once the principle of QTLs mapping had been established, statistical tools were developed to identify and localise QTLs involved in quantitative variation of the trait under study on the linkage map.

IV.4. Identification and localisation of QTLs: statistical approaches to QTLs mapping

The simplest method consists of analysing the data using one marker at a time. This does not require a complete molecular marker linkage map. This single marker approach assumes that QTLs can only occur exactly at the marker locus tested. At each marker, a t-test is carried out to test significance of contrasts of marker class means. In a backcross population issued from crosses between two inbred lines, contrast between marker class means is:

$$(\mu_{M1/M1} - \mu_{M2/M2})/2 = a(1-2c)$$

where a is the additive effect, c is the recombination frequency between the marker and the QTL, $\mu_{M1/M1}$ and $\mu_{M2/M2}$ being the phenotypic mean of marker classes M1M1 (homozygous for allele M1 at the marker locus) and M2M2 (homozygous for allele M2 at the marker locus). The parameter tested is then composite, including additive effect and recombination frequency. As the two parameters of interest (a and c) cannot be estimated separately, this approach does not allow a distinction between tight linkage to a QTL with small effect and loose linkage to a QTL with large effect. It follows that this approach leads to an incorrect localisation of the QTL and an

underestimation of additive effect if the QTL does not lie at the marker locus (Martinez and Curnow, 1992). Another single marker approach has been developed by Kearsey and Hyne (1994), relying on regression analysis. This method still relies on the linear relation $\delta_i=a(1-2c)$, with δ_i being half the difference between the means at the ith marker. At the true position of the QTL, there should be a linear regression of δ_i on $(1-2c)$ with a slope a passing through the origin. This approach is achieved for a set of linked markers, with the QTL placed initially at the first marker. The regression analysis is then carried out at regular intervals, every 2cM, along the chromosome. The most likely position for the QTL is the location at which the residual sum of squares of the regression is minimised. This process can be repeated for all sets of markers in order to determine locations and effects of all QTLs influencing the trait.

The most commonly used method of QTLs mapping is called interval mapping analysis (or flanker-marker analysis) and was defined by Lander and Botstein (1989). Pairs of linked markers are analysed simultaneously. This approach is based on a likelihood function, taking into account the observed data (number of individuals and their phenotypes in each marker class) and the parameters to be estimated (recombination frequency, means and variances of the QTL genotypes). The test of significance is based on a log likelihood ratio, or LOD score ($\log_{10}L/L_0$ L being the observed likelihood assuming one QTL in the interval, and L_0 the likelihood under the null hypothesis of no QTL segregating). An iterative computer program enables identification of the maximum likelihood estimates (MLEs) of the unknown parameters that maximise the observed likelihood function L given the data. The LOD score is computed over several positions along the marker interval; the most probable position of the QTL is the one that maximises the LOD score. MLEs and standard error estimates of the MLEs are taken from the interval with the highest LOD score. Typical LOD score values for presence of a significant QTL lies between 2 and 3 (Lander and Botstein, 1989).

Darvasi et al. (1993) determined the effect of several factors on the power of QTL mapping and on the standard error estimates (SEE) of MLEs of gene effect and map location. The resolving power of any experiment is determined by population size and the effect of the QTL. Marker spacing inferior to 10-20cM provides no additional gain of power on SEE. The same authors compared single marker analysis and interval mapping and their results concur with those

of Tanksley (1993). Indeed, the maximum advantage of interval analysis over single marker analysis is realised when linked markers are around 20-35cM apart. When marker density is higher (<20cM), both methods give similar results. When marker loci are further apart (>35cM), neither method is highly efficient at detecting QTLs. Even if the use of an interval mapping method does not increase power of the analysis compared to the use of single method, the accuracy of parameter estimation is increased. A further difficulty is that estimates of locations and effects of QTLs are asymptotically unbiased if the assumption that there is, at most, one QTL on a chromosome is true. If the assumptions of the model are violated then the parameter estimates may be biased. Therefore Lander and Botstein (1989) proposed to extend their method to the analysis of multiple markers, at one time, to map multiple QTLs. Nevertheless this multidimensional search (complex and requiring larger sample size) is still biased because the number of QTLs on a chromosome is unknown and the assumptions of the model may not be fulfilled.

Multiple regression analysis (Haley and Knott, 1992; Martinez and Curnow, 1992) is based on least squares methods for performing flanking marker analyses for QTL detection. Similar to the interval mapping approach, it uses information from 2 linked flanking markers, and the presence of a QTL in this interval is tested. The test consists of fitting the one-QTL linear model to find the parameter estimates that minimise the residual sums of squares (RSS); if the presence of a QTL in that interval is accepted, the position associated to the minimum RSS gives the most likely position of a QTL as well as the best estimates of its effects. Interval mapping analysis (maximum likelihood approach) and multiple regression analysis give very similar estimates of QTLs locations and effects (Haley and Knott, 1992). These two approaches are indeed very similar as they consist of testing the presence of a QTL in a interval flanked by two markers. It is simply the statistical basis underlying them which is different. One advantage of the multi-regression approach is that it can be performed using general computer statistical software, contrary to the interval mapping approach which requires a specific software (Mapmaker/QTL). Moreover, multiple regression analysis can be extended to test for the presence of two or more QTLs, to test interactions between QTLs, or to analyse threshold traits (Haley and Knott, 1992). In addition, Martinez and Curnow (1992) highlighted the effect of QTLs in neighbouring regions of the chromosome, leading to erroneous locations and estimated effects of a QTL when

information from flanking markers alone is included in the analysis. To overcome this problem and discriminate between the presence of one or two QTLs, these authors suggest a bivariate regression mapping with three markers. The presence of one or two QTLs is detected by looking at the RSS surface generated. The main limitation of these multidimensional searches is the larger sample size required for a sufficient power detection. Jansen (1993) developed an alternative approach, still based on multiple regression analysis, which allows a more efficient detection and a more accurate mapping of QTLs, by reducing the interference of neighbouring QTLs. Their procedure first selects markers closely located to QTLs by a multiple regression of the quantitative trait on all markers. Multiple QTLs on a chromosome are therefore detected. These markers are then used as cofactors during the regression of phenotype on genotype and expected to neutralise some of the effects of other QTLs.

Interference of neighbouring QTLs on the estimation accuracy of location and gene effect of a QTL in a interval flanked by two markers has been identified previously (Lander and Botstein, 1989; Haley and Knott, 1992; Martinez and Curnow, 1992; Jansen, 1993). However, attempts to overcome this problem were based on multiple QTLs models, using information of multiple markers, which can be complex and powerful only with very large sample size. These approaches are therefore often difficult to use. A very interesting approach has been developed by Zeng (1993). It consists of creating an interval test, i.e. to test whether there is a QTL in the interval delimited by two markers i and i+1. This test is achieved by conditioning on two markers bracketing the interval (i-1 and i+2). The test statistic constructed is unaffected by QTLs located outside the defined interval. This approach reduces a multidimensional search problem (multiple QTLs models approach) to a one-dimensional search problem because the presence of a QTL is tested in one interval at one time. Composite interval mapping (Zeng, 1994) is an interval test procedure based on the methodology developed by Zeng (1993) which combines interval mapping with multiple regression analysis to make a full use of marker data. Test for the presence of a QTL in the interval (i, i+1) is achieved by a likelihood ratio test statistic ($LR=-2\ln L_0/L_1$), where L_0 is the likelihood function under the null hypothesis of no QTL segregating and L_1 the likelihood function under the alternative hypothesis of presence of a QTL in the interval. Maximum likelihood estimates of location and effect of the QTL are obtained by an iterative algorithm as the test is performed at any position along the genome. In this way a

likelihood profile is obtained along the entire chromosome, allowing the mapping of all QTLs determining the trait under study. This approach leads to a powerful detection of QTLs and to accuracy in mapping QTLs and estimating their effects. Indeed, composite interval mapping focuses on one region at a time (defined by two flanking markers), and makes the test of presence of a QTL in that interval independent of other regions by conditioning on all other markers (these markers are fitted in the one-QTL model, controlling genetic background). However, with limited data, power and precision of mapping can be gained by adopting a semi-composite interval mapping: only a few markers that have been selected by stepwise regression are fitted in the model to control the genetic background. Another extension of the composite interval mapping method, developed by Jiang and Zeng (1995), consists of analysing multiple traits simultaneously. This joint analysis can improve the power of detecting QTLs and the precision of parameter estimation. Moreover, and more importantly, this approach allows the testing of biological hypotheses such as pleiotropy, QTL by environment interaction, or pleiotropy versus close linkage.

Finally, several more recent studies have used Bayesian methods to fit multiple-QTL models, more precisely using Markov Chain Monte Carlo (MCMC) algorithms for parameter estimation and hypothesis testing (Jansen, 1996; Satagopan et al., 1996; Uimari et al., 1996; Sillanpaa and Arjas, 1998). This approach seems more suitable to the analysis of outbred populations. It can be implemented by classical statistical software. However, Sillanpaa and Arjas (1998) developed a software programme, Multimapper, that can deal with inbred lines or outbred populations (http://www.rni.helsinki.fi/~mjs).

IV.5. Identification and location of QTLs: experimental design

QTLs mapping relies on linkage disequilibrium between marker alleles and QTL alleles. Therefore, experimental designs that maximise this linkage disequilibrium can greatly improve mapping efficiency and accuracy. There are two main ways of creating such populations in linkage disequilibrium. Firstly, crosses between inbred lines, each fixed for different alleles, produce F1 individuals heterozygous at all loci that display maximum disequilibrium. Secondly, high linkage disequilibrium results from sampling a collection of relatives.

Several experimental designs can be implemented to maximise linkage disequilibrium. A classical approach involves crosses between two inbred lines, to form either backcross or F2 populations. Most of statistical methods developed are suitable for this kind of experimental design. Selective genotyping consists of scoring individuals for the trait and then genotyping only individuals selected from the tails of the character distribution. Linkage disequilibrium is likely to be maximal between individuals presenting extreme phenotypes. Darvasi and Soller (1992) showed that selective genotyping can reduce the number of individuals genotyped by seven fold for a given power of detecting QTL. Doubled haploid lines (DHLs) are produced by treating chemically gametes from parents to restore diploidy. The cross between two doubled haploid individuals (completely homozygous) can lead to a maximum linkage disequilibrium in the progeny. It mimics the case of inbred lines in species where such lines cannot be produced.

Several methods can increase statistical power of QTL mapping by reducing the within-marker-classes standard deviation. Indeed, the power to detect a difference in mean between two marker genotypes depends on the difference scaled by this standard deviation (Falconer and Mackay, 1996). Progeny testing consists of asexually replicating each genotype. Mean values of these replicated progeny will be taken into account instead of individual values, leading to more accurate estimates of phenotypic values. Recombinant inbred lines (RILs) are produced from an F1, following by multiple cycles of auto-fertilisation or multiple generations of brother-sister mating. There is no within-line genetic variance whereas the between-lines variance is large because each RIL represents a different multilocus genotype.

Another efficient approach to map individual QTLs is to use nearly isogenic lines (NILs). These lines are inbred lines containing one or more fragments introgressed from a donor parent. Such lines are constructed by crossing a donor parent to a recurrent parent (from an inbred line). The resulting offspring is then backcrossed to the recurrent parent for several generations, so that a certain proportion of the donor parent is retained in each inbred line. Each NIL retains different portions of the genome. Comparing phenotypic values with marker genotypes across a collection of NILs can permit relatively precise mapping of QTLs.

When no linkage map is available, a very efficient method can be employed to detect QTLs and identify markers that are linked to these QTLs: bulked segregant analysis (BSA). This method, developed by Michelmore et al. (1991), relies on the distribution of marker alleles among groups of individuals that present extreme phenotypic values. Pooled DNA is used; band intensities between the two groups of individuals are compared to detect differences between groups. Initially, this method was used to saturate a target region with markers because it is a useful tool for the construction of a high precision linkage map. However when applied to QTL mapping, this approach is expected to detect only QTLs associated with high gene effect (Lynch and Walsh, 1998). Marker alleles in linkage disequilibrium with QTL alleles are expected to have a non-random distribution across groups of individuals. The group of individuals with the lowest phenotypic value will be enriched for one or several markers alleles whereas groups of individuals with the highest phenotypic value will be enriched for other marker alleles. Unlinked markers are expected to have a random distribution across groups.

QTL mapping in outbred lines uses the same principle as for inbred lines in that it relies on marker-trait associations in segregating populations, more often F2 populations as this design is the most powerful. However, QTL mapping is considerably more difficult in outbred populations mainly because not all parents will be informative. Informative parents are those that exhibit segregation at both markers and QTL alleles. One method to increase the power of the analysis is to select as parents those individuals that exhibit different marker genotypes such as to maximise linkage disequilibrium in their progeny. The advantage of controlled matings when working with outbred populations has been discussed by Tanksley (1993).

IV.6. Software of QTL mapping

Several software programmes have been developed to map QTLs (Table 6, Manly and Olson, 1999).

Table 6. Current mapping software. From Manly and Olson (1999).

Software	Functions	Method	Designs	Authors	URL
Mapmaker/QTL	SIM, CIM	ML	BC, F2	Lincoln et al. (1992)	ftp://genome.wi.mit.edu/pub/mapmaker3
QTL Cartographer	SIM, CIM	ML	diverse	Basten et al. (1997)	http://statgen.ncsu.edu/qtlcart/
Map Manager QT	SIM, CIM	LS	BC, F2, RI	Manly and Olson (1999)	http://mcbio.med.buffalo.edu/mapmgr.html
MapQTL ® 5	SIM, CIM	ML	BC, F2, RI, DH	Van Ooijen (2004)	http://www.cpro.dlo.nl/cbw/
PLABQTL	SIM, CIM	LS	BC, F2	Utz and Melchinger (1996)	http://www.uni-hohenheim.de/~ipspwww/soft.html
MQTL	SIM, sCIM	LS	BC, DH, RI	Tinker and Mather (1995)	ftp://gnome.agrenv.mcgill.ca/pub/genetics/software/MQTL
Multimapper	CIM	Bayesian	BC, F2	Sillanpaa and Arjas (1998)	http://www.rni.helsinki.fi/~mjs
QTL Express	SIM	LS	Inbred, outbred pop	Seaton et al. (2002)	http://qtl.cap.ed.ac.uk

SIM: simple interval mapping; CIM: composite interval mapping; sCIM: simplified composite interval mapping.
ML: maximum likelihood; LS: least-squares regression
BC: backcross; F2: F2 population (cross between F1 individuals); RI: recombinant inbred lines; DH: doubled haploid lines.

IV.7. QTLs mapping in fish aquaculture species

Most of the literature found concerned fish species, and particularly rainbow trout, for which numerous studies reported the finding of QTLs. QTLs were detected for traits of economic importance in fisheries: spawning time, maturation timing, disease resistance, growth, lower or upper temperature (UTT) resistance. The QTLs detected explained some part of the phenotypic variance of the trait, ranging from 4 to 34% individually. Most of the time, several QTLs were found for a single trait, confirming the polygenic inheritance of those traits.

The most widely used experimental designs were backcrosses or F2 intercrosses of hybrids between two strains or species with strong phenotypic divergence, in order to maximise the likelihood of identifying QTLs. Backcrosses of phenotypically divergent strains have been used widely to detect QTLs for UTT (Jackson et al., 1998; Danzmann et al., 1999; Somorjai et al., 2003), spawning time (Sakamoto et al., 1999) or IHNV (Infectious Hematopoietic Necrosis Virus) resistance (Ozaki et al., 2001; Rodriguez et al., 2004; Khoo et al., 2004).

A powerful method for the identification and mapping of QTLs was employed by Palti et al. (1999): Bulk Segregant Analysis (BSA). According to these authors, requirements for a successful BSA include that: i) the pooled DNA samples compared originate from a single cross between two individuals that differ strongly in the trait (resistant x susceptible parents). ii) a large number of segregating individuals is used in each group, such as to eliminate variation not associated with the trait of interest. To increase efficiency of BSA approach, Palti et al. (1999) applied a gene-introgression approach (inter-species crosses) combined with BSA: they detected three candidate fragments, two associated with susceptibility to IHNV and one to resistance. Moreover, Nakamura et al. (2001) reported that BSA combined with the AFLP technique is a very powerful approach for identifying markers tightly linked to genes underlying monogenic traits. Groups were made by pooling equal amounts of DNA from 15 albino and 15 normal individuals. Identification of candidate markers (64 primer pairs tested) was followed by screening these candidates in each of the progeny and the location of the gene responsible for autosomal dominant albinism was identified. The BSA/AFLP approach also represents a useful tool for increasing marker density in a target region and identifying specific genes. Another good example of the utilisation of BSA for detecting marker-phenotypic trait associations was reported by Lee et al. (2003) for the identification of a sex-determining region in tilapia. It is important to notice that BSA approach is not only restricted to mapping of monogenes and could be used to map QTLs with moderate to high effect. This approach was successfully used for mapping QTLs in rainbow trout (Palti et al., 1999; Rodriguez et al., 2004) and in Kuruma prawn (Li et al., 2006a).

One way of controlling against false discovery of QTLs consists of using two independent experiments. The first step is a genome scan in the first family and then only the markers that

showed associations in the first family are screened in a second family, for confirmation (Cnaani et al., 2004).

Disease resistance is an ideal candidate for the application of marker-assisted selection (MAS), due to its economic importance. This trait is of particular interest for shellfisheries and we can imagine that in the future MAS for disease resistance could be applied in some shellfish species. The effectiveness of MAS in a selection process based on challenge with the pathogen depends on the proportion of additive genetic variance explained by markers and on heritability of the disease trait. Several studies have reported the location of QTLs for disease resistance in rainbow trout, based on the classical approach for QTL mapping using interval mapping, the ANOVA-based approach, or BSA (Palti et al., 1999; Ozaki et al., 2001; Rodriguez et al., 2004; Khoo et al., 2004). Recently, Moen et al. (2004a) have showed that QTLs affecting disease resistance can be detected in species for which genetic maps are not available. The approach employed represents a shortcut compared with other QTL mapping strategies currently used. This multistage QTL testing strategy is potentially more powerful than conventional genome-wide scans where thresholds must be set very high to account for the testing of a large number of genetic markers. According to these authors, two designs can be used to map QTL for disease resistance in fish:

- Selection of high and low ranking individuals followed by testing for homogeneity of genotypic classes in the two groups. Parents are ranked according to the average number of surviving progeny. Markers are genotyped in the two extremes of the distribution.
- Transmission disequilibrium test (TDT) for QTL mapping. This relies on the assumption that affected offspring would inherit marker alleles in a larger proportion than expected because of linkage with QTL affecting resistance to a disease. Test compares if the frequency of alleles M and m is the same among dead progeny in a challenge experiment. An advantage is that genotyping efforts are focused on the affected offspring (Moen et al., 2004a).

The identification and location of QTLs for several traits and in several species highlights the potential for marker assisted selection in breeding programmes in fisheries and probably shellfisheries, in the future. Sakamoto et al. (1999) report that identification of a large number of DNA markers linked to QTL controlling traits of economic significance will contribute to the

application of DNA marker-assisted selection (MAS) in aquaculture breeding. However, Danzmann et al. (1999) show evidence of epistatic interactions between UTT QTL alleles and genomic background in rainbow trout. Therefore, effectiveness of MAS may be limited without prior knowledge of the performance characteristics of QTL alleles in different genomic backgrounds. Moreover, Perry et al. (2001) tried to confirm QTLs detected in experimental crosses between two divergent strains in outbred populations and some trait-markers associations were retained in these outbred populations (therefore confirming the potential for MAS). The important role of genetic background was stated to explain differences between outbred populations and divergent lines. Moreover, Ozaki et al. (2001) highlighted the importance of investigating the role of epistasis in determining the expression of IPN (Infectious Pancreatic Necrosis) disease resistance/susceptibility in rainbow trout. Perry et al. (2003) showed the importance of epistasis influencing trait expression, in particular the strong role of the sex chromosomes in regulating the effects of autosomal genes for growth and stress resistance traits in rainbow trout. Therefore, once QTLs have been located, it is important to assess the role of epistasis and genomic background on the effects of these QTLs, in order to test the feasibility of MAS.

Most of the studies reported the location of QTLs in a quite large area, spanning 10 to 60 cM. Therefore, it is not clear if a QTL might be made up of several genes with more effects, some acting in opposing direction. Before applying MAS, it is therefore important to achieve fine QTL mapping and restrict the region of interest to a more narrow area to increase the likelihood of successful MAS. MAS has a huge potential in aquaculture breeding programme, but so far no successfully applied MAS has been reported in fish or shellfish species. A further step will be to go from the QTL down to the specific genes, to better understand the genetic underlying expression of quantitative traits. BAC (Bacterial Artificial Chromosome) libraries will facilitate positional cloning of QTLs, and already BAC libraries have been constructed in several aquaculture species, e.g. Japanese flounder (Katagiri et al., 2000), rainbow trout, carp, tilapia (Katagiri et al., 2001) and the Pacific oyster (Hedgecock et al., 2005).

Table 7. Review of QTLs (Quantitative Trait Loci) detected in fish and shellfish aquaculture species. V_A: variance, LGs: linkage groups, BC: backcross, FS: full-sibs, DH: double haploids, BSA: bulked segregant analysis, UTT: upper temperature tolerance, IHNV: Infectious Hematopoietic Necrosis Virus, IPNV: Infectious Pancreatic Necrosis Virus, ISA: Infectious Salmon Anemia, TDT: transmission disequilibrium test, IM: interval mapping, CIM: composite interval mapping, ANOVA: analysis of variance, ANCOVA: analysis of covariance.

Species	QTLs detected	Experimental design	Statistical analysis	References
Oncorhynchus mykiss Rainbow trout	Upper temperature tolerance 2 QTLs detected V_A explained: 22% Markers: 36 allozyme loci, 10 RAPDs, 19 microsatellites	Two strains used, one with low tolerance to high temperature (L), the other with high tolerance (H). Three BC families issued from cross between an hybrid male (HxL) and two females H and one female L.	QTL analysis performed separately on maternal and paternal alleles. Associations of individual marker loci with UTT were tested by a two-way ANOVA with an interaction term (genotype by experiment) and by a two-way ANCOVA (fork length used as a covariate). Sequential Bonferroni correction.	Jackson et al. (1998)
	Upper temperature tolerance 3 QTLs detected	The same than Jackson et al. (1998)	Two-way ANCOVA. Test for epistasis between QTLs alleles and genomic background	Danzmann et al. (1999)
	Spawning time 5 chromosomal regions in 5LGs (repeatable effects) Markers: 54 microsatellites	Two strains used, a spring spawning (S) and a fall spawning (F). One BC family issued from cross between an hybrid male (SxF) and a female F.	QTL analysis performed separately on maternal and paternal alleles. One-way ANOVA to test associations between marker alleles and spawning time.	Sakamoto et al. (1999)
	IHNV resistance 2 markers associated with susceptibility and one with resistance. DNA fingerprinting	16 BC families, issued from 4x4 factorial mating design between an hybrid male (rainbow x cutthroat) and a rainbow female. Cutthroat trout less susceptible to IHNV. 3 challenge replicates for each BC family.	ANOVA performed to test the sire and dam effects, and the interaction sire*dam on survival. BSA analysis to detect markers linked to IHNV resistance. DNA fingerprints of two pools (15 first mortalities/15 survivors) compared to parents.	Palti et al. (1999)
	• Time to hatch 2 QTLs detected, V_A: 24.6% • Embryonic length 2 QTLs detected, V_A: 22.6% • Embryonic weight 2 QTLs detected, V_A: 26.2% Markers: 219 AFLPs, 2 microsatellites	2 doubled haploid families issued from a cross between two homozygous clonal lines showing divergent hatching time. DH families come from F1 males (SWxOSU). Line SW exhibits an accelerated development rate.	Composite interval mapping. Software used: QTL cartographer (model 6, window size of 5cM; 5 markers as cofactors).	Robison et al. (2001)

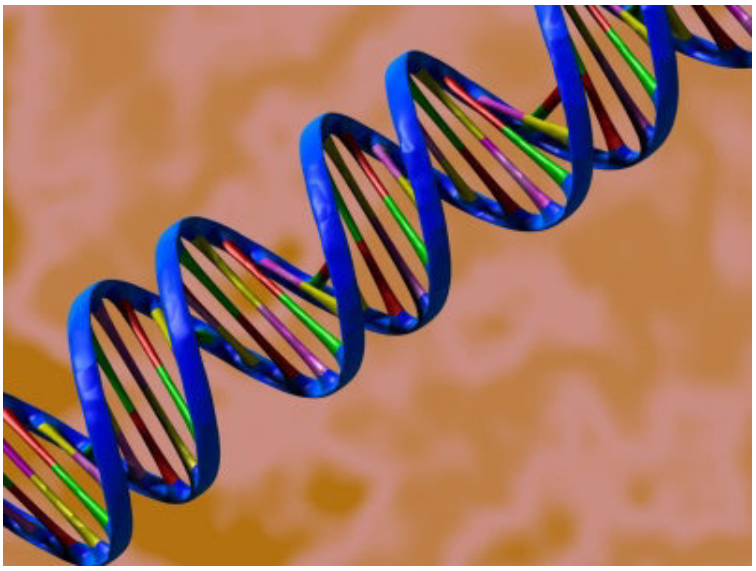
	IPNV resistance 2 QTLs detected V_A explained: 27-34% Markers: 121 microsatellites	Two phenotypically divergent strains, one resistant to IPN virus (R), the other susceptible (S). One BC family issued from the cross of an hybrid male (RxS) and a female S.	First, identification of potential markers linked to QTLs by a χ^2 distribution (52 BC progeny analysed, 26 that died and 26 that survived at 51 microsatellite loci). Then, potential markers scored in the entire BC progeny (54 dead, 46 survivors). Finally, interval mapping (software: Map Manager QT) to locate and estimate gene effects.	Ozaki et al. (2001)
	• Spawning time 4 significant QTLs and 2 suggestive (on 6 LGs) V_A explained: 8-64% individually • Body weight 3 significant QTLs (on 3 LGs) V_A explained: 12-25% individually Markers: 201 microsatellites	Two strains used, a spring spawning (S) and a fall spawning (F). One BC family issued from cross between an hybrid male (SxF) and a female F. 90 progeny. Continuity of the study of Sakamoto *et al.* (1999).	QTL analysis performed separately on maternal and paternal alleles. Interval mapping. Software: MultiQTL. Two models used, one assuming similar trait variance, the other unequal variances. Permutation test.	O'Malley et al. (2003)
Oncorhynchus mykiss Rainbow trout	Body mass 5 QTLs V_A explained: 4-7% Markers: microsatellites located on the same LGs previously associated with body mass.	4 paternal half-sib families. Two cultured strains were used: SV (autumn spawning) and RS (spring spawning). 2 PHS families come from crosses between 2 SV males and 15 SV females; the 2 other PHS families come from crosses between 2 RS males and 20 RS females.	Multiple regression using least squares means (SAS). Model includes a fixed effect for the dam, and one for the tank. Significance levels determined using permutation-derived thresholds	Martyniuk et al. (2003)
	IHNV resistance 18 AFLPs and 6 microsatellites linked to QTLs (3 LGs mostly involved).	7 BC families issued from cross of several F1 families (steelhead x rainbow) with rainbow trouts.	QTL analysis performed separately on maternal and paternal alleles. BSA analysis: independent segregation of the marker and the QTL is tested by a χ^2 analysis between two groups (dead and survivors).	Rodriguez et al. (2004)
	7 meristic traits (counts of vertebrae, scales, rays…). At least one QTL detected for each of six traits. Markers: 20 AFLPs primer sets	Inter-strain cross, between two clonal lines. DH progeny were produced from the sperm of F1 hybrid (OSUxCW).	Composite interval mapping. Software used: QTL cartographer (window size of 2cM; 5 most significant markers in the stepwise linear regression are used as background control markers). Significance thresholds determined by permutation test.	Nichols et al. (2004)

Oncorhynchus mykiss Rainbow trout	IHNV resistance 4 closely linked microsatellites associated with IHN resistance (on the same LG), identifying one chromosomal region implied. Markers: 96 microsatellites	5 BC families issued from crosses between two Japanese strains (one resistant to IHNV, the other susceptible). 92 progeny per BC.	Binary code for tracking alleles coming from the resistant strain or the susceptible one. Associations between allele markers and phenotype (dead/survival) tested by Pearson's χ^2 test. Significance for $p<0.05$.	Khoo et al. (2004)
	Pyloric caeca number 3 QTLs identified, located on 3 LGs. Span 66 cM. 16 genetic markers (15 AFLPs, 1 microsatellite) are associated with those 3 QTLs. V_A explained: 13-19% individually	OSU x HC double haploid population (n=54). These 2 strains were used because they exhibit high difference in pyloric caeca number. Sperm of a F1 male was used for androgenesis.	Composite interval mapping. Software used: QTL cartographer (model 6, window size of 5cM). Markers to be used as cofactors were identified by linear regression.	Zimmerman et al. (2005)
Salvelinus alpinus Arctic charr	Upper temperature tolerance 2 significant QTLs, and 7 suggestive Markers: 40 microsatellites	Crosses between two strains divergent for temperature tolerance. 4 F1 families (♂x♀, ♀x♂) and one BC family (♂F1x♀ strain less resistant) obtained. Microsatellite loci were chosen from known association with QTL in rainbow trout (knowledge of homeologs).	Selective genotyping in two families. QTL analysis performed separately on maternal and paternal alleles. One locus analysed at a time. Comparison of alleles classes using Kaplan Meier product limit measure (non parametric test). Correction of time of death by body weight.	Somorjai et al. (2003)
Salmo salar Atlantic salmon	ISA resistance 2 putative QTLs associated with resistance in one FS family. Markers: 64 AFLP primer sets	One paternal half-sib family comprising two FS families of 40 and 39 offspring.	First, transmission disequilibrium test performed on affected fish. Then, Mendelian segregation test achieved for each marker significant in the TDT. Finally, all TDT-significant markers following Mendelian segregation are used in the survival analysis (χ^2 test comparing groups of individuals having the band, and those who don't). Permutation test to account of multiple testing.	Moen et al. (2004a)
Cyprinus carpio L. Common carp	Cold tolerance 4 putative markers associated with cold tolerance, one of which was mapped.	Interspecies cross. 100 F2 offspring challenged for overwintering. These offspring derived from crosses between F1 hybrids (common carp *C. carpio* x Boshi carp *C. pelligrini*). Local common carp highly resistant to low temperatures; Boshi carp highly sensitive to low temperatures.	2 pools of DNA were compared, the survivors and the dead fish. Pools compared using RAPD-PCR. Aim: identifying bands amplified only from DNA samples of the cold tolerance grandparent and the F2 survivors. Then mapping of the putative markers in a haploid mapping panel (used to build a linkage map).	Sun and Liang (2004)

Oreochromis sp. Tilapia	• Body weight / Standard length One QTL influencing both traits. • Cold tolerance One QTL located on the same LG. Additive and dominance effects estimated. Markers: 20 microsatellites	Two F2 hybrid families (mating of F1 hybrids issued from inter-species crosses between *O. mossambicus* and *O. aureus*. 60 and 114 FS offspring per family.	First, one-way ANOVA for each maker individually (20 microsatellites) to test associations between DNA markers and cold tolerance and body weight (family of 60 FS). Then, regression interval mapping (QTL Express software) using data of 6 microsatellites present on the same LG (family of 114 FS).	Cnaani et al. (2003)
	14 markers on 9 LGs were associated with stress response traits (innate immunity response, blood parameters). V_A explained: 9-13% Markers: 40 microsatellites and 2 genes	Two F2 hybrid families (mating of F1 hybrids issued from inter-species crosses between *O. mossambicus* and *O. aureus*). 79 and 114 FS offspring per family.	Tests for associations between DNA markers and phenotypic traits by one-way ANOVA for each marker-trait combination. First, scan of the 40 microsatellites in the first family; then, scan of 9 microsatellites showing associations with stress response in the second family.	Cnaani et al. (2004)
	Cold tolerance 2 putative QTLs located in LG 23 Markers used: AFLPs, 54 microsatellites	Four-way crosses between four species of tilapia: *O. aureus, O. mossambicus*, *O. niloticus*, *Sarotherodon galilaeus*. Two 4WC families ((On x Sg) x (Oa x Om)) of 54 (family 1) and 44 offspring (family 2).	First, single marker analysis in the family 1 to test associations between all informative markers and cold tolerance or body weight (t-tests with permutation tests). Then, t-tests achieved in family 2 for significant marker-trait associations. Finally, interval mapping (QTL Express software) to estimate position of the putative QTLs. Allele contributions for the 2 parents analysed separately. One-QTL model.	Moen et al. (2004c)
Penaeus japonicus Kuruma prawn	Growth traits: weight, total length, carapace length. No significant QTL in female. One QTL (both IM and CIM) and one suggestive QTL (CIM only) detected in male. V_A explained: 23% Markers: 54 AFLP primer sets	F2 full-sib population of 443 individuals, issued from cross between growth-selected lines. HH high selected line for weight, LL low selected line. F2 population issued from reciprocal mating of 4 HH with 4 LL, and one F1 HL maternal and one F1 LH paternal.	Selective genotyping strategy on a full-sib F2 intercross design. Interval mapping (Zmapqtl model 3) and composite interval mapping (model 6, window size of 5 cM, 2 markers as cofactors). Software: QTL cartographer. BSA-derived markers included to increase number of markers in the QTL region.	Li et al. (2006a)

CHAPTER 2

Molecular markers



Deoxyribonucleic acid (DNA) is a nucleic acid that contains the genetic instructions for the development and function of living organisms. DNA is a long polymer of simple units called nucleotides, which are held together by a backbone made of sugars and phosphate groups. This backbone carries four types of molecules called bases (adenine, cytosine, guanine, and thymine) and it is the sequence of these four bases that encodes information. The structure of double helix of the DNA was figured out by Crick and Watson in 1953. The major function of DNA is to encode the sequence of amino acid residues in proteins, using the genetic code. In eukaryotes such as animals and plants, DNA is stored inside the cell nucleus, while in prokaryotes such as bacteria, the DNA is in the cell cytoplasm.

A gene is a DNA fragment that contains genetic information and can be defined as a unit of inheritance that is transmitted from parents to offspring. It includes segments of the DNA that are transcribed into a functional RNA product, as well as enhancers, regulatory regions and introns. The place where a particular gene resides on a chromosome is called a locus (plural, loci). Variant forms of the gene are termed alleles. The genome represents all hereditary information (haploid DNA complement). In most eukaryotes, each individual has two copies of a given locus, they are diploid. The alleles at the various loci represent the genotypes. At a single locus, diploid individuals will either be homozygous if the two copies are identical or heterozygous if the two copies are different.

Molecular markers are simply DNA fragments (genetic information) or their molecular representations such as RNA and proteins. Molecular markers exhibit several features. Firstly, variability of molecular markers is called polymorphism: several individuals will exhibit different forms (or alleles) at a particular locus. Polymorphism of molecular markers is defined by a finite number of allelic states, in contrast to phenotypic traits that exhibit a continuous variation. Some molecular markers will be characterised by low polymorphism, others by high polymorphism, depending on the number of alleles present in a population. Secondly, molecular markers can be either defined as specific or non specific. Specific markers are amplified by polymerase chain reaction (PCR) by one primer pair that recognise specifically the fragment to be amplified (e.g. microsatellites). Non specific markers are PCR-amplified by universal primers or randomly-chosen primers that can reveal polymorphism at several loci following a single PCR

amplification (e.g. Randomly Amplified Polymorphic DNAs: RAPDs or Amplified Fragment Length Polymorphisms: AFLPs). Thirdly, molecular markers can be codominant or dominant. In genetics, dominance relationship refers to how the alleles for a single locus interact to produce a phenotype. If the marker is codominant, both alleles of a diploid individual are detectable such that heterozygous and homozygous genotypes can be unambiguously identified. However, a dominant marker will be analysed in terms of phenotypes, i.e presence or absence of the marker. The phenotype "absence of the marker" will be associated to the genotype "homozygous for the recessive allele", whereas the "presence of the marker" will be ambiguous in terms of genotypes, either heterozygous or homozygous for the dominant allele. Fourthly, molecular markers will either exhibit a biparental heredity (nuclear DNA markers where the two copies of a given allele are inherited through both parents, one coming from the father, the other from the mother) or uniparental heredity (e.g mitochondrial DNA markers that exhibit maternal transmission). Finally, molecular markers are either neutral or under selection.

Allozymes correspond to the analysis of protein polymorphisms. Polymorphism corresponds to a modification of the net charge of a protein, revealed by migration in an electrical field (electrophoresis). These markers are codominant, mono-locus, reliable, low-cost and exhibit biparental heredity. First developed in the 1960s, allozyme data dominated the genetic literature until the 1990s when they were replaced by reliable DNA-based markers. Identification and scoring of DNA-based polymorphisms relies either on the hybridisation of a probe or on the exponential amplification of the locus of interest by Polymerase Chain Reaction (PCR) (Saiki et al., 1988). PCR amplifies a specific fragment of DNA by using short oligonucleotides that are complementary to the DNA sequences at each end of the fragment. Millions of copies of the fragment are made using *Taq* polymerase and nucleotide units that double the number of fragments each cycle during repeated temperature cycling (Saiki et al., 1988). PCR-based polymorphisms can be revealed after electrophoresis by either a variation of fragment length, of restriction site (sites of DNA that are cut by restriction enzymes), or on the sequences of DNA itself (nucleotide changes).

In this chapter, the focus will be on the two types of DNA-based markers, microsatellites and AFLPs (Amplified Fragment Length Polymorphism), that were used during this study.

I- MICROSATELLITE MARKERS

I.1. Definition and general features

Variable Numbers of Tandem Repeats (VNTRs) or Simple Sequence Repeats (SSR) are a class of highly informative and widely dispersed genetic markers in DNA (Shriver et al., 1993). Different classes can be defined, considering the repeat unit size:

- Microsatellites are short fragments of DNA made up of sequences repeated in tandem of arrays of 2 to 6 bp (Figure 7).

- Minisatellites are composed of repeats units of 15-70 bp.

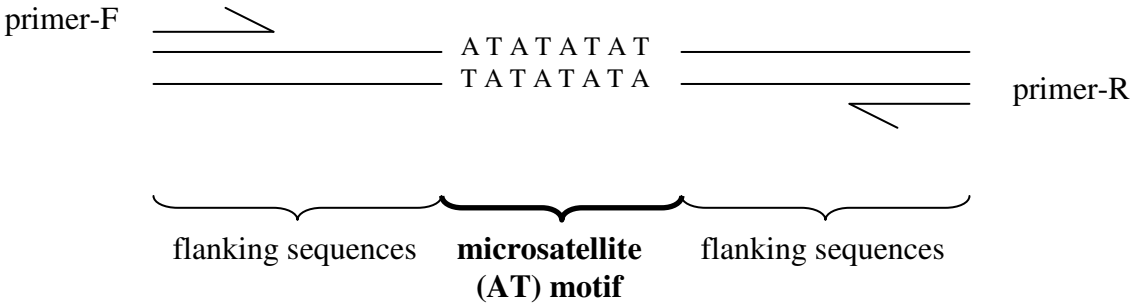


Figure 7. Microsatellite defined as a motif repeated in tandem. Flanking sequences on which primer pairs anneal for revelation of polymorphism are shown.

Microsatellites can be specifically amplified by Polymerase Chain Reaction (Saiki et al., 1988). These sequences are embedded within unique sequences (the flanking sequences) which allow their amplification once these sequences are known and once specific pairs of primers have been designed (Dallas, 1992). Three groups of microsatellites were initially defined by Weber (1990): perfect, imperfect and compound. Perfect repeats consist of sequences with no interruption in the motif of repeats (e.g. $(CA)_{20}$), imperfect repeats consist of sequences with one or more interruptions in the repeats (e.g. $(TG)_{16}TT(TG)_{11}$), and compound repeat sequences consist of perfect or imperfect repeats adjacent to another simple repeat sequence (e.g. $(GT)_{14}(GA)_{16}$ (Wintero et al., 1992)). The most abundant are repeats with a mono or dinucleotide motif (Rassmann et al., 1991). Morgan and Rogers (2001) affirm that the most commonly occurring microsatellites are poly(A) and poly(T) but that these regions are unstable. So dinucleotide repeats such as $(CA)_n$ are the most commonly used in genetics.

The major features of the microsatellite markers justify their wide use. They are codominant, highly variable, are expected to be neutral and are inherited according to Mendelian rules. Most studies make the assumption that these markers are neutral. However, the functional role of microsatellites has been reviewed by Li et al. (2002) and showed that the repeat family is implicated in a variety of functions including chromatin organisation (chromosomal organisation, DNA structure, centromere and telomere organisation), regulation of gene activity, recombination, DNA replication or cell cycle.

I.2. Abundance (number of copies) and distribution

Microsatellites are numerous in all prokaryotic and eukaryotic genomes that have been examined and present both in coding and noncoding regions of the genome. The abundance of these blocks of repetitive units has been well studied in the human genome. The $(AC)_n$ motifs occur every 30 kb according to Litt and Luty (1989); the $(CA)_n$ blocks, copy number estimated as 50000-100000, appear every 30-60 kb (Weber and May, 1989) as well as the blocks $(TG)_n$ which are also present in about 50000 copies. More recently, Toth et al. (2000) analysed microsatellite frequency in a wide range of taxa, from fungi to primates and confirmed that microsatellites are very abundant, with an occurrence of one microsatellite every 2 to 11 kb depending on the taxon considered.

Microsatellites have been reported by several authors to be distributed at random throughout most eukaryotic genomes (Hearne et al., 1992) and in particular in the human genome (Weissenbach et al., 1992). Wong et al. (1990) studied the distribution of microsatellite markers in 10 species of vertebrates, from fish to human and demonstrated that these markers are distributed throughout the chromosomal arms. However, more and more studies showed that the distribution of microsatellites is not homogeneous within a single genome and also across taxa in terms of the absolute numbers of microsatellites and repeat motif preference (Hancock, 1999; Toth et al., 2000). Different repeat motif will be more abundant according to different regions of the genome, intergenic regions, introns or exons. A statement in one taxon will not hold in another taxon.

I.3. Finding microsatellites

The traditional method to find microsatellite markers is based on the screening of genomic libraries by simple repeat oligonucleotide probes (Weber and May, 1989; Rassmann et al., 1991; Queller et al., 1993). The different steps of this protocol involves: (i) genomic DNA extraction and digestion by restriction enzymes; (ii) selection of the fragments comprised in the size range 400-800bp; (iii) digestion of the cloning vector (plasmid or phage) by a restriction enzyme and ligation of the DNA fragments; (iv) transformation of the vector in competent bacteria; (v) screening of these genomic libraries by labelled di-, tri- or tetra-oligonucleotide probes; (vi) identification of the positive clones; (vii) DNA extraction, sequencing of the insert fragments to define the primers. This approach has been successfully used by Weissenbach et al. (1992) to isolate a large number of CA repeats in the human genome.

The PIMA approach (PCR isolation of microsatellite arrays) begins with a step of RAPD amplification, following by cloning of the fragments into a vector, transformation and PCR screening of clones to identify clones containing microsatellites (Lunt et al., 1999).

Several enrichment methods have been developed in order to reduce the time and efficiency of microsatellites isolation (reviewed by Zane et al., 2002). These alternative methods result in the production of libraries enriched in microsatellite loci. Primer extension reaction is an enrichment protocol in which a single strand DNA library is primed with repeat-specific oligonucleotides (Ostrander et al., 1992; Paetkau, 1999). The most commonly used enrichment protocol is based on a digestion-adaptor ligation step followed by hybridisation with one or several repeat-containing probes, PCR amplification and cloning of the microsatellite-enriched DNA. Several “selective hybridisation” protocols have been developed (Karagyozov et al., 1993; Armour et al., 1994; Kandpal et al., 1994; Kijas et al., 1994; Edwards et al., 1996). Finally, Zane et al. (2002) proposed another enrichment protocol: FIASCO (Fast Isolation by AFLP of Sequences COntaining repeats) based on MseI digestion of the DNA and ligation with adaptors, followed by PCR amplification with MseI-N primers, hybridisation of the DNA with a biotinylated $(AC)_{17}$ probe and selectively capturing DNA molecules hybridised with probes.

The isolation and development of microsatellite markers is a very time consuming task. But once developed, their use is easy and routine. Owing to the fact that their development can take some time, an interesting feature of the microsatellites reported in numerous studies is the possibility of cross-priming amplifications between related species. Most of these studies report the success of cross-amplification for closely related species only (Moore et al., 1991; Primmer et al., 1996; D'Amato et al., 1999). Moore et al. (1991) report the conservation of dinucleotide microsatellites among mammalian genomes and therefore the cross-amplification in sheep of microsatellite markers initially developed in cattle. Moreover, the study of D'Amato et al. (1999) indicates the possibility of successfully using the microsatellite markers developed for the hake *Macruronus magellanicus* in other gadoid species. Finally, Primmer et al. (1996) tested cross-species amplification of sets of primers developed from the swallow and the pied flycatcher on 48 other bird species. These authors found a significant and negative relationship between the performance of cross-amplification and the evolutionary distance between the original species and the tested species. This negative correlation can be explained by mutations in the flanking regions of the microsatellites. The accumulation of mutations in the flanking sequences will not permit the annealing of the primers and there will be no amplification of the markers in species which are more distantly related.

I.4. Polymorphism

The length of microsatellites varies considerably according to the number of repeats, which is roughly comprised between 15 and 30 (Litt and Luty, 1989; Weber and May, 1989). The fact that the polymorphism observed is a result of the variation of the number of repeats of the motif has been demonstrated by Rassmann et al. (1991). This considerable polymorphism can be visualised (thanks to the use of labelled fluorescent primers) by electrophoresis on a denaturing polyacrylamide gel. This kind of gel allows the discrimination of bands which differ only by 1bp, improving considerably the sensitivity of the technique.

The vast majority of microsatellite alleles are <50 repeat units in length. Garza et al. (1995) showed that constraints may occur to limit the size of the microsatellite loci. The mechanism which controls the allele length at such loci is not well identified: gene conversion, biased mutation (towards the lower length size) or selection could be responsible for this

observation. Therefore, polymorphism at microsatellite is high but within a certain range of size length.

The possibility of scoring multiple loci on the same gel has been reported by several authors (Weber and May, 1989; Queller et al., 1993). Recently, automated sequencing machines have improved the multiple-scoring. Indeed, the primers of each marker can be labelled with different fluorescent molecules; therefore more than one marker can be analysed at the same time even if the size ranges of these markers overlap. There is therefore no more need to select microsatellites that have size ranges that are different enough from one another to permit the visualisation of more than one microsatellite on the same gel.

When genotyping microsatellite markers, additional bands, less intense than the major pair of bands, can often be seen (Weber and May, 1989). This is called “stutter”. These extra, less intense fragments are generated during the amplification reaction and are not a sign of somatic mosaicism. These bands reflect truncated products generated during the PCR by mispairing of the template DNA and the newly synthesized strand during elongation (Hearne et al., 1992). More often they represent a major problem for interpretation and scoring. However, these additional bands can be in some instance useful to determine the size of the allele since they differ in size by one or more repeat unit, representing a size scale.

Of the three types of microsatellites, perfect, imperfect and compound, Rassmann et al. (1991) suggest that loci with composite patterns should be avoided, preferring the use of perfect dinucleotide motifs because the scoring of such polymorphism is easier and less likely to be subject to PCR artefacts.

I.5. Origin of polymorphism- Mutations

The high polymorphism of microsatellites implies that they have relatively high mutation rates. The mutation rate can be estimated using three methods: direct counting of mutations on pedigrees; indirect estimation from linkage data or comparison of observed and theoretical values (Jarne and Lagoda, 1996). Numerous studies have tried to estimate the mutation rates of these markers, revealing variability in the estimation according to the loci studied and the number of

repeats of the motif or the species studied. The mutation rate of microsatellites in the human genome is estimated at between $5x10^{-4}$ and 10^{-5} by Hearne et al. (1992), at $1.2x10^{-3}$ on average for 28 loci by Weber and Wong (1993) whereas Weissenbach et al. (1992) estimated this rate close to 10^{-2}. In recombinant inbred strains of mouse, Dallas (1992) estimated the upper limit of microsatellite mutation rates between 10^{-2} and 10^{-4}, similar to the interval found in the human genome. According to the mutation rates found, microsatellites are suitable for genetic mapping, as mutations are unlikely to be observed in the pedigrees analysed which represent fewer than 1000 meioses (Dallas, 1992; Hearne et al., 1992). Therefore, these markers provide a good compromise in the trade-off opposing the variability and the mutation rates since they are associated with a high polymorphism and relatively acceptable mutation rates (Queller et al., 1993). Different studies demonstrate that the mutation rates can greatly vary according to the size of the repeated motif involved. Indeed, according to Weber and Wong (1993), the average mutation rate in human for the tetranucleotide loci was nearly four times higher than the average rate for the dinucleotide one. The study of Viard et al. (1996) in a freshwater snail agreed with Weber and Wong (1993): the lower variability of the dinucleotide loci compared with the tetranucleotide is linked to a lower mutation rate. On the contrary, Chakraborty et al. (1997) estimated the mutation rates of the non-disease-causing dinucleotide loci in humans to be 1.5 to 2 times higher than the tetranucleotides; for the disease-causing trinucleotide loci, the mutation rates are about 3.9 to 6.9 times larger than for the tetranucleotides. Moreover, the mutation rate varies according to the type of microsatellite sequence. Garza et al. (1995) showed that imperfections in the repeated region may cause a decrease in the mutation rate of microsatellite loci.

The variation of the number of repeats at microsatellite loci has been thought to arise by two main mechanisms reviewed by Levinson and Gutman (1987): replication slippage (Kornberg, 1980; Schlotterer and Tautz, 1992) and unequal meiotic exchanges (Hamada and Kakunaga, 1982; Jeffreys et al., 1985; Treco and Arnheim, 1986). Replication slippage implies the dissociation of the two strands followed by a re-hybridisation of the strands associated with a mispairing of complementary bases. This mispairing at a microsatellite locus results either in a diminution or in an increase of the number of the repeated motif. For VNTRs, slippage appears to be the predominant mode of new allele formation (Valdes et al., 1993). According to the same

authors, in vitro experiments lead by Schlotterer and Tautz (1992) indicate that the rate of slippage is dependent on the size of the repeat unit (greatest for dinucleotides) and on its sequence (slowest for GC-rich repeats). Moreover, Lai and Sun (2003) estimated that the slippage mutation rate increases exponentially as the number of repeat units increases and that expansion occurs more frequently for short microsatellites and contraction occurs more frequently for long microsatellites; this observation is congruent with the size constraints on the repeated motifs which clearly does not allow an indefinite increase in size. Unequal meiotic exchanges occur during recombination when tandemly repeated sequences are aligned out of register. Hamada and Kakunaga (1982) report the possibility that the $(TG)_n$ sequences, highly dispersed in the human genome, adopt a Z-DNA conformation and then play a role in recombination and/or control of gene expression. Moreover, the analysis of tetrads showed that these same repeated sequences, when inserted in the genome of *Saccharomyces cerevisiae*, enhance reciprocal meiotic recombination by a factor of seven and cause multiple recombination events to occur within a relatively small region of the yeast chromosome (Treco and Arnheim, 1986). Therefore, these sequences, susceptible to adopt a Z-DNA conformation, may represent a signal to initiate genetic exchange. Finally, Jeffreys et al. (1985) demonstrated that minisatellite sequences are actively implicating in unequal exchange at a high rate and suggested that these sequences could act as a recombination signal, creating a kind of hotspot for crossing-over. Levinson and Gutman (1987) argue that slippage replication accounts mainly for the initial expansion of microsatellite motifs and that the tandem repeats continue their further expansion by unequal crossing-over. Origin of polymorphism of microsatellites is still under debate (Chambers and MacAvoy, 2000; Ellegren, 2004), but replication slippage seems the most likely mechanism (Zane et al., 2002).

Despite the wide use of microsatellites, their mutational mechanism is still not plainly known. Different theoretical mutation models have been proposed to explain the mode of mutation of these markers, the Infinite Allele Model (IAM) or the Stepwise Mutation Model (SMM). The infinite allele model IAM (Kimura and Crow, 1964) stipulates that each mutation gives rise to a new allele not found in the population, resulting in an infinite number of allelic states. The stepwise mutation model (SMM) was first introduced by Ota and Kimura (1973). It stipulates that one mutation event results in the changes of one unit of what we observed (net charge for allozymes, number of repeats for microsatellites). The number of alleles is then

limited. Expansions and contractions are supposed to be equally likely (Whittaker et al., 2003). This model was first developed to account for the mutational mechanism of allozyme markers and then explain the allelic distribution observed for these markers. More recently, this model has been shown not to be applicable to allozymes but to potentially be much more relevant to the evolution of microsatellites.

Shriver et al. (1993) linked theoretical mutation models to the underlying molecular mechanism of mutation: unequal exchanges during recombination can result in the potential for a very high number of alleles and therefore could be the molecular mechanism underlying the IAM model; on the contrary, slippage replication usually leads to changes in length of one repeat (increase or deletion of one unit repeat), compatible with the one-step SMM model. Simulations were used to assess the validity of a mutation model to a data set in such a way as to find the most suitable model and so better understand the origin of the high polymorphism of microsatellites (Shriver et al., 1993; Valdes et al., 1993). As a conclusion, the SMM model seems more suitable for microsatellite loci (implying as most likely mechanism of mutation replication slippage), whereas the IAM could be more likely for the minisatellites (with unequal exchanges as an underlying mode of mutation). However, it seems difficult to draw general conclusions. Indeed, depending on the size of the repeat unit, the microsatellite loci fit more or less the one-step SMM (Shriver et al., 1993). In addition the mutational mechanism could vary according to the types of microsatellite.

More recently, more complex models of mutation have been proposed for microsatellites. Di Rienzo et al. (1994) analysed frequency distributions of allele size, at the population level, in human populations (Sardinian, sub-Saharan African and Egyptian). These authors introduced a two-phase mutation model, which assumes that most mutations change allele size by one repeat unit only (one-step phase: increase or decrease of one repeat unit) but the rest change allele size by several repeat units (multistep phase). Garza et al. (1995) introduced a bias mutation model to take into account constraints on allele size.

I.6. Applications

Sunnucks (2000) reports that microsatellites have many potential applications. First, the general features of the microsatellite markers, high polymorphism, neutrality, codominance and Mendelian inheritance, make them ideal for parentage analysis (Queller et al., 1993). Microsatellites have been used to assess paternity in wild populations. The first type of method, based on an exclusion procedure in which the multilocus genotypes of the offspring are compared locus per locus with those of the potential progenitors (Ellstrand, 1984), have been used in wild chimpanzees (Morin et al., 1994) and in toque macaques (Keane et al., 1997). Another approach, based on the calculation of likelihood ratio and the analysis of the LOD score (=Likelihood Of Odds), has also been used to infer the paternity in red deer (Marshall et al., 1998; Slate et al., 2000) and harbour seal (Coltman et al., 1998) populations as well as in a cohort of Atlantic cod larvae (Herbinger et al., 1997). Valenzuela (2000) clustered hatchlings of side-neck turtles in function of their pairwise relatedness coefficients, to identify full-sibs groups. Microsatellites have also been used to infer the parentage in experimental crosses. For example, individual parentage has been determined in Atlantic salmon using compatibility analysis, by comparison of alleles locus per locus, from each offspring with those of the potential parents used for the crosses (O'Reilly et al., 1998). Boudry et al. (2002) used parentage data to highlight high variance in reproductive success in the cupped oyster *Crassostrea gigas*.

Moreover, microsatellites are markers of choice for high density genome mapping. The large number of the $(CA)_n$ repeats in the human genome make their use suitable in genetic mapping (Weber and May, 1989). Indeed, Weissenbach et al. (1992) constructed a linkage map of the human genome with an average resolution of 5cM based on the segregation of 814 polymorphic $(CA)_n$ loci.

Microsatellites have been used to estimate intra-population and inter-population variability in a range of plant and animal species, such as a tree species (Collevatti et al., 2001), the Chinook salmon (Kinnison et al., 2002), the European flat oyster (Launey et al., 2002), the eastern cupped oyster (Rose et al., 2005) or the Pacific cupped oyster (Li et al., 2006c). Moreover, population parameters of crucial importance can be estimated, such as the effective size of the population *Ne*, the selfing fertilization rate *S* or the migration rate *m* (e.g. Viard et al.,

1996). The estimation of these parameters allows a better understanding of the forces which act to maintain the genetic diversity and structure of the populations studied. Microsatellites have also been used successfully in phylogenetic construction and species identification (Primmer et al., 1996).

A more specific, narrow use is the specific larval identification from mixed plankton samples (Morgan and Rogers, 2001), allowing new applications for population genetics at earlier stages of life.

I.7. Cautions in the use of microsatellite markers

Microsatellites may be subject to size homoplasy (Weber and Wong, 1993) and some loci can be associated with the segregation of null alleles (Callen et al., 1993).

A null allele is an allele that does not produce an identifiable product on the gel. The most common cause is mutations (single-base pair substitution, insertion, deletion) occurring in the flanking sequences that prevent the binding of a primer, resulting in the non-amplification of the allele during the PCR. Null alleles are not restricted to microsatellites: they have been previously demonstrated at allozyme loci and reported by Foltz (1986) as a possible cause of heterozygote deficiencies in bivalves. Null alleles have been shown at human microsatellites (Callen et al., 1993; Phillips et al., 1993). Moreover, McGoldrick et al. (2000) reported that 16 microsatellite loci of the 24 studied in *C. gigas* present null alleles and that they are inherited and thus stable. Therefore, null alleles are frequent in Pacific oyster and more generally should be taken into account when studying bivalve populations. Launey et al. (2002) suggest that the heterozygote deficiency at one locus observed by them in populations of *O. edulis* is mainly due to the segregation of null alleles.

In population genetic studies, it is therefore very useful to estimate the frequency of null alleles to identify the part of the heterozygote deficit attributable to the presence of the null allele in the population. Two methods of calculation of the frequency of null alleles have been developed. First, Chakraborty et al. (1992) developed a method of estimation of null allele frequencies (r) based on the apparent deficiency of heterozygotes. Secondly, Brookfield (1996)

estimated the null allele frequency taking into account the existence of non amplified individuals, considered as homozygotes for the null allele. These two methods give minimal and maximal estimates respectively for the null allele frequencies.

Caution should be exercised in parentage analysis. The segregation of a null allele in a family can result in mismatches between the putative parents and the offspring, resulting in the exclusion of one potential parent which was in fact the real parent. Therefore, in most parentage studies, a certain degree of mismatches is allowed where offspring and putative parent appear to be different homozygotes (Pemberton et al., 1995), permitting the transmission of a null allele in the offspring (Figure 8).

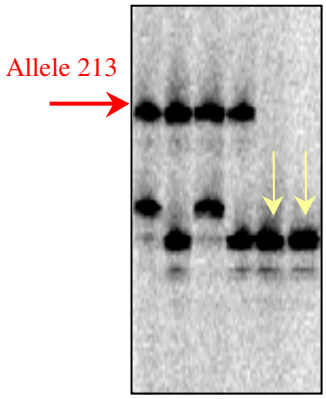


Figure 8. Example of mismatch between the genotypes of the known parent and their offspring. Each column represents one offspring. Alleles of the known parent are showed by red arrows. The two yellow arrows highlight two offspring whose genotype is not compatible with the parent's genotype (homozygote for the allele 213). This can be interpreted by the segregation of a null allele in the family.

Finally, Callen et al. (1993) reported that in gene mapping, the presence of a segregating null allele will not corrupt the linkage data but simply results in loss of informativeness. Moreover, McGoldrick et al. (2000) affirm that null alleles can be accommodated in gene mapping as long as the number of markers used are at least between 50 and 100.

The phenomenon of size homoplasy is the occurrence of two different alleles that are of the same size due to different mutational events which occur independently. Estoup et al. (1995b) studied in detail the occurrence of homoplasy in two bee species, *Apis mellifera* and *Bombus terrestris*. They estimated the amount of size homoplasy at different levels: intra-population and inter-population for closely related and more distant populations by sequencing electromorphs of

two loci and then comparing the sequences. These authors showed that size homoplasy occurs frequently between distantly related subspecies but is less common at a lower population level. Caution must be taken in population genetics since unidentified homoplasy leads to a clear underestimation of genetic distances. Moreover, interrupted microsatellites are less subject to homoplasy; therefore, their informativeness makes them suitable to investigate population differentiation and evolutionary relationship between relatively distant populations (Estoup et al., 1995a).

II- AFLP (AMPLIFIED FRAGMENT LENGTH POLYMORPHISM) MARKERS

AFLP is a DNA fingerprinting technique which detects genomic DNA restriction fragments by means of PCR amplification. This technique was first reported and described in detail by Vos et al. (1995).

II.1. Principle

The AFLP methodology consists of the following steps. First, pure genomic DNA of high molecular weight is digested by two different restriction enzymes, one a frequent cutter (4-bp recognition site) and the other a relatively infrequent cutter (6-bp sequence). The two most commonly used enzymes are MseI and EcoRI respectively. But, the choice of these cutters depends on the organism studied. For example, Ajmone-Marsan et al. (1997) report that the use of the couple TaqI/EcoRI is more suitable for the analysis of mammalian genomes. The restriction digestion results in three categories of fragments which differ by length: short fragments generated by the frequent cutter only; long fragments generated by the infrequent cutter only; and intermediate fragments which result from the digestion by both enzymes.

The second step consists of the ligation of oligonucleotide adapters to the restriction-digested DNA fragments. The double-strand adapters consist of a core sequence and an enzyme-specific sequence which allows the annealing of the adapters to the recognition-site of a specific restriction enzyme. This step is very important because it will permit the PCR amplification of the fragments using primers that are complementary to the core sequence of the adapters. The

sequences of the adapters are chosen so that the restriction sites are not recreated after the ligation.

The third stage consists of the amplification of a subset of all the restriction fragments. This step is called the preselective amplification. The primers used consist of a core sequence, an enzyme specific sequence and an additional single selective nucleotide. The stringent PCR conditions used allow the amplification only of the fragments which match exactly with the primers sequences, i.e. the fragments which possess the nucleotide complementary to the selective primer. This additional nucleotide at the 3' end of the primers allows the amplification of only a subset of the total fragments generated during the restriction digest/ligation steps. Thus, this step enables a reduction in the complexity of the band patterns. The primers are unlabelled because no visualization of the PCR products is needed at this stage.

Finally, a selective amplification is made on the subset of pre-amplified fragments. The structure of the primers is the same as the preselective primers except that this step uses 3 or more selective nucleotides which allows further reduction in complexity of the band patterns obtained. The more the number of selective nucleotides is increased, the more the subset of fragments amplified is reduced. For complex genomes, primers with three selective nucleotides appear to give the best results (Vos et al., 1995). Only one of the two selective primers is labelled; the labelled one is used to completion whereas the unlabelled one is in excess. During the amplification, only the fragments having the two different cleavage sites (fragments of intermediate size, <600bp) are subsequently amplified (Vos et al., 1995; Sharbel, 1999). Finally, the PCR products can be visualized by denaturing polyacrylamide gel electrophoresis or by automated analysis on a capillary sequencing machine (Figure 9).

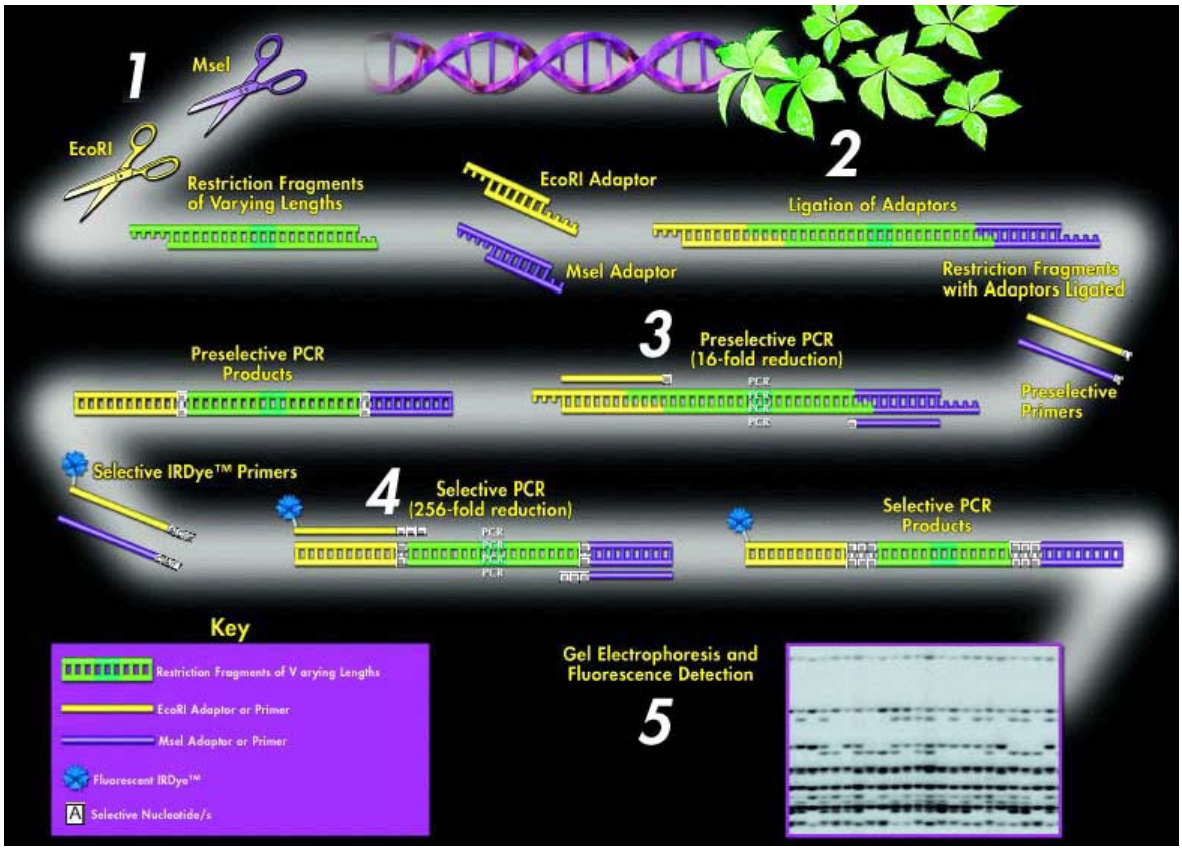


Figure 9. AFLP methodology (http://www.licor.com/bio/SAGAMX/AFLP12.jsp). After extraction of the DNA from any source, the AFLP methodology is performed and consists of several steps. 1- restriction digestion; 2- ligation of adaptors; 3- preselective PCR; 4- selective PCR; 5- electrophoresis for detecting polymorphism.

II.2. General features

Mendelian inheritance of AFLP markers has been shown in several studies (Maughan et al., 1996; Ajmone-Marsan et al., 1997; Maheswaran et al., 1997).

The AFLP technique is repeatable but the quality of the genomic DNA used is important. If the quality is poor, the restriction digests will not be complete, leading to a non-reproducibility in the AFLP banding pattern (Matthes et al., 1998). So, great caution must be accorded to the purity of the DNA: different methods for purifying the DNA are available. The repeatability of the AFLP technique was reported to be very high. Vos et al. (1995) tested the banding patterns obtained with an increased number of selective nucleotides during the selective amplification. They showed that the bands obtained with 4 selective nucleotides are present in the pattern obtained with 3 selective nucleotides; only the number of bands differs between the two experiments. In other words, the addition of an extra selective nucleotide always resulted in a

fingerprint which was a subset of the original fingerprint. These authors also demonstrated that the AFLP procedure is insensitive to variations in the template DNA concentration above the minimum threshold of 2.5pg although Sharbel (1999) recommends equalizing the DNA concentrations across the samples studied. Moreover, AFLP profiles have been shown to be highly reproducible among seven different laboratories (Jones et al., 1998). Finally, numerous studies report the reproducibility of the AFLP technique by comparing the banding pattern obtained for the same individual across independent DNA extractions and template preparations (e.g. Ajmone-Marsan et al., 1997; Escaravage et al., 1998). The high reliability of this technique is mainly due to the stringent PCR conditions for the annealing of the primers and can be opposed to the low stringency of the simpler RAPD (Randomly Amplified Polymorphic DNAs) method.

Data obtained from AFLP markers are dominant. The basic assumption is that each band corresponds to a locus at which two alleles segregate, determined either by the presence (dominant allele) or the absence of the band. Individuals scoring a band may be homozygous or heterozygous for the allele determining the band. Nevertheless, some studies report the possibility of scoring the AFLPs markers as codominant markers (Ajmone-Marsan et al., 1997; Sharbel, 1999), by allowing identification of the heterozygotes. This can be done by comparing band intensities or peak heights between individuals. But care should be taken, because one has to be sure that differences in band intensities do not come from other causes such as different DNA concentrations, or differences between PCR amplification within, or between batches.

Finally, polymorphisms are detected as the presence or absence of bands, each band representing a locus. This variation can be caused in three main ways. Firstly, there can be differences in restriction sites. Secondly, there can be mutations around the restriction sites, preventing the annealing of the primers during the amplification cycles. These mutations may cause mismatch of the fragment or the core sequence or the selective nucleotides in the primers. Thirdly insertions or deletions may occur within the amplified restriction fragment (Matthes et al., 1998). This third type of polymorphism can cause errors in interpretation because the two alleles both produce bands, but they are of different size. This polymorphism at one locus may be scored as two separate loci.

The size range for the fragments obtained is typically between 50 and 450bp. The size of each band can be determined using size standards. The amount of polymorphism obtained with each array (one pair of primers) depends on the organism studied. Usually, to increase the informativeness of the AFLP technique, different combinations of primers pairs can be used. In this way, even if the PIC (polymorphic informative content) is low for each locus, the great number of loci increases considerably the degree of information from this technique. Krauss (1999) reports that different AFLP patterns can be obtained with template DNA extracted from distinct organs of the same plant due to differences in DNA methylation between organs (Donini et al., 1997). In most of the studies, when the AFLP banding pattern is analysed as dominant markers, the recessive allele frequencies (q) are calculated from the frequency of the recessive phenotype q^2 (band absent) for each polymorphic locus assuming HW equilibrium.

II.3. Applications

The first potential application of the AFLP technique is in the field of genetic diversity and genetic relationships analysis. The genetic diversity, using the AFLP data, has been assessed within and between populations. For example, Travis et al. (1996) used 220 AFLP markers to assess the genetic diversity within 3 populations of plant *Astragalus cremnophylax* and show a genetic differentiation among these three populations. Moreover, the AFLP technique has been successfully used to study the genetic diversity within and among species. For example, Hill et al. (1996) determined phylogenetic relationships among different species of *Lactuca spp*. which are congruent with the known taxonomic relationships for these taxa. Finally, Hardy (2003) showed that AFLP markers can be as efficient as microsatellites to assess relatedness between individuals and characterize the spatial genetic structure of populations. Moreover, AFLP markers represent an efficient alternative to microsatellites in population assignment studies which aim is to identify the population source of an individual (Campbell et al., 2003).

The second field of applications is in parentage analysis. Most paternity or parentage studies are conducted using microsatellites as genetic markers. However, Gerber et al. (2000) demonstrated, through analysis on oak tree data comparing the results obtained with AFLP and microsatellite markers, that parentage analysis can be performed with a good statistical confidence with AFLP markers even if these dominant markers present lower exclusion

probabilities than the codominant one. This result is very important and shows that parentage analysis can be performed relatively easily and quickly, with a minimal cost, in natural populations in species for which codominant markers have yet to be developed. It is an important feature that no prior knowledge of the genome is required for AFLP analysis. The following studies illustrate the use of AFLP markers in parentage analysis. Questiau et al. (1999) analysed the segregation of 81 polymorphic AFLP markers (obtained with the combination of 3 primer pairs) in 36 families of a passerine bird species, the bluethroat: both parents and offspring were genotyped. This study demonstrated the possibility of extra-pair mating in 63.8% of the broods, indicating fertilization by other male (or males) than the social male (who takes care of the nest). Secondly, the utility of AFLP technique for paternity analysis in natural populations has been demonstrated by Krauss (1999) on a population of Proteaceae plant *Persoonia mollis*. Paternity was successfully assigned for 96% of the seeds using 3 primers pairs and the paternity of the remaining seeds was assigned by adding two sets of primers. At this time, the main limitation for the success of paternity analysis is the dominant feature of the AFLP markers. The probability of exclusion (which determines the power of exclusion of the potential fathers by the set of markers used) is lower for dominant markers (Questiau et al., 1999; Gerber et al., 2000). This would be expected because the heterozygotes cannot be distinguished from the dominant homozygotes, reducing the possibility of discriminating two individuals based on their genotype data. Therefore, the possible identification of heterozygotes at AFLP loci through the quantification of the band intensity or peak height will improve the utility and increase the power of exclusion of AFLP markers for paternity analysis (Krauss, 1999). Nevertheless, the lack of resolution between the heterozygotes and homozygotes is compensated by the high number of polymorphic loci generated after a single PCR-assay: the probabilities of exclusion obtained with the AFLP markers are indeed a little lower than those for microsatellites but remain high enough to achieve paternity exclusion in good conditions.

Finally, AFLPs can be used in genetic mapping and linkage analysis. The AFLP markers can be used to construct high density maps of genomes or genome segments because of their ability to detect a great number of fragment polymorphisms in a single PCR-array (Vos et al., 1995). Indeed, no prior information on the genomic sequences is required; more than 50 loci per single PCR-array can generally be generated, these loci being distributed throughout the genome.

So these markers are very efficient for the construction of genetic linkage maps (Otsen et al., 1996). AFLP markers have been used to increase the resolution of a number of maps. For example, Otsen et al. (1996) placed AFLP markers on a pre-existing genetic map of the rat based on 12 biochemical, 2 immunological, 1 RFLP and 55 SSCP markers. The results of this study show the potential of AFLP markers for mapping of QTLs since they found a correlation between two closely linked AFLP markers and mean arterial pressure. Since then, AFLP markers have been used in numerous studies for linkage analysis and detection of QTLs. Most of the time, these markers are combined with microsatellite markers to improve the power of resolution of such genetic maps (See Chapter 1 for the full review).

III- MATERIAL AND METHODS

III.1. DNA extraction protocol

The necessity of a high-quality DNA has been reported in several studies. Therefore, three different protocols were used to obtain high-molecular weight and high quality DNA.

First, a standard phenol-chloroform DNA extraction method as described in Sambrook et al. (1989) was achieved in 10 samples of flat oyster. For each oyster, around 1mm² of gill was digested in 400 µl of lysis buffer (100 mM NaCl, 10 mM Tris-HCl pH=8.0, 25 mM EDTA pH=8.0 and 0.5% SDS) and 5 µl proteinase K (10 $mg.ml^{-1}$) at 55°C overnight. DNA extraction was then performed by adding 400 µl of phenol-chloroform-isoamyl alcohol (25:24:1) and centrifugation for 10 minutes at 10000 rpm. The DNA (supernatant) was precipitated by adding 400 µl of 100% ethanol (-20°C), waiting at least 2 hours at –20°C and centrifugating at 12000 rpm for 10 minutes. Finally, the pellets were washed by adding 400 µl of 70% ethanol (-20°C) and centrifugating at 8000 rpm for 5 minutes, then dried and re-suspended in 50 µl autoclaved TE (10 mM Tris-HCl pH=8.0, 1 mM EDTA pH=8.0).

The second method was described in Grewe et al. (1993). For each oyster, around 1mm² of gill was digested in 700 µl of 2% CTAB buffer (100 mM Tris-HCl pH=8.0, 1.4 M NaCl, 20 mM EDTA, 2% CTAB) and 10 µl proteinase K (10 $mg.ml^{-1}$) at 60°C for 2 hours. DNA was then

extracted by adding successively 600 µl of chloroform-isoamyl alcohol (25:1), 600 µl of phenol-chloroform-isoamyl alcohol (25:24:1) and 600 µl of chloroform-isoamyl alcohol (25:1) : each step was followed by centrifugation at 10000 rpm for 10 minutes, the supernatant being used for the next step. The aquaeous phase was then mixed in an equal volume of isopropanol and samples were placed at –20°C for 2 hours for the precipitation of the DNA. All samples were then centrifuged at 12000 rpm for 10 minutes. DNA pellets were finally washed twice with 500 µl 70% ethanol (-20°C) followed by centrifugation at 8000 rpm for 5 minutes. DNA was dried and re-suspended in TE.

The third method consisted of a chloroform extraction followed by purification with the Wizard® DNA Clean-Up System (Promega). For each oyster, around 1-2 mm² was digested in 400 µl of 2% CTAB buffer (100 mM Tris-HCl pH=8.0, 1.4 M NaCl, 20 mM EDTA, 2% CTAB) and 20 µl proteinase K (10 $mg.ml^{-1}$) at 60°C overnight. Then, 2 successive chloroform extractions were performed by adding 400 µl of chloroform-isoamyl alcohol (25:1) followed by centrifugation at 13200 rpm for 5 minutes. The upper layer (about 300 µl) was then cleaned with the Promega kit. One Wizard® Minicolumn and one syringe are used for each sample to avoid contaminations between samples. Each sample was mixed by inversion with 1 ml of Wizard® DNA Clean-Up Resin. The resin containing the bound DNA was inserted into a 2 ml syringe barrel attached to a minicolumn. The syringe plunger was then inserted and pushed slowly to allow the retention of DNA onto the minicolumn. Then the DNA attached to the minicolumn was washed by pushing 2 ml of isopropanol 80% through the minicolumn with the plunger. The minicolumn was then centrifuged at 13200 rpm for 2 min to dry the resin and 50 µl of pre-warmed TE (65-70°C) was added: after one minute of waiting to allow DNA to detach from the minicolumn, a final centrifugation at 13200 rpm for 20 seconds allows the elution of the bound DNA.

Whatever the DNA extraction method used, DNA concentration was measured for each sample by spectrophotometry (Ependorf Bio Photometer) and the ratio of optic density DO at 260 nm / DO at 280 nm. The 260/280 nm ratio allows us to check that DNA samples are not contaminated by proteins. A “good quality” DNA sample will have a ratio DO 260 / DO 280 between 1.7 and 1.8. Each sample was adjusted to give a final concentration of 100 $\mu g.ml^{-1}$.

The three DNA extraction methods described were carried out on gill samples from the same 10 samples of flat oysters. It appeared that the quality of the DNA obtained was far higher when using the Wizard® DNA Clean-Up System (Promega): on agarose gel, a band of high molecular weight was obtained, with a 260/280 nm ratio always between 1.7 and 1.8 (Figure 10). On the contrary, the two other methods produced degraded DNA, visible as a strong smear on agarose gel and with DO ratios sometimes above 2.0. See Figure 11 for visual comparison between the classical phenol chloroform method and the DNA Wizard® Clean Up System one. No RNase step was perfomed in any of the 3 protocols. It can be seen Figure 11 that the phenol-chloroform DNA extraction protocol led to the extraction of RNA as well as DNA.

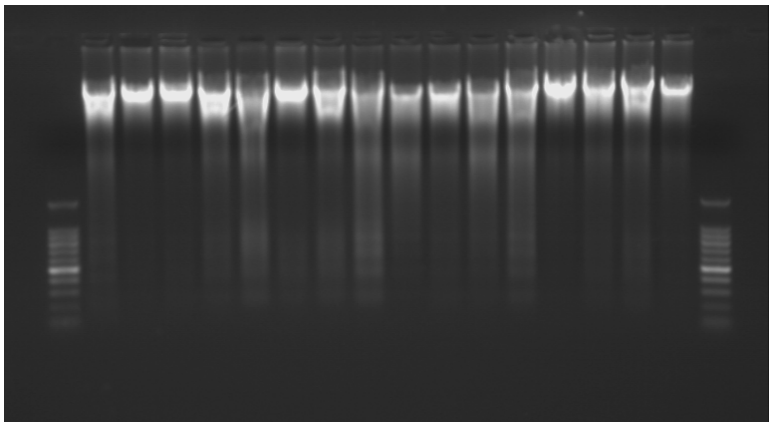


Figure 10. Visualisation on a 2% agarose gel of a high molecular weight DNA extracted with the Wizard® DNA Clean-Up System (Promega).

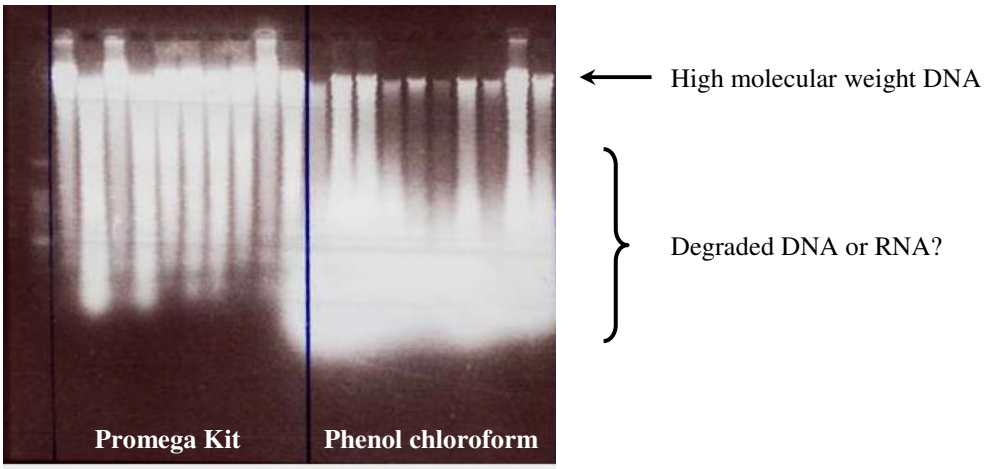


Figure 11. Visual comparison between 2 DNA extraction protocols, chloroform DNA Clean Up System (Promega) and phenol chloroform.

Further developmental trials compared the DNA extraction efficiency from gill or adductor muscle. DNA concentrations were between 60 and 180 $\mu g.ml^{-1}$ with adductor muscle and between 200 and 800 $\mu g.ml^{-1}$ with gill tissue. Due to the high quantity of DNA required for our experiments, gill tissue was chosen for achieving DNA extraction.

In conclusion, the DNA extraction method chosen was a chloroform extraction from gill tissue followed by purification with the Wizard® DNA Clean-Up System (Promega).

III.2. Optimisation of microsatellite markers in *Ostrea edulis* and *Mytilus edulis*

Electrophoresis and data collection was carried out on an ABI 3100-Avant (Applied Biosystems). Two main fluorescent dyes were used: FAM and HEX (Eurogentec). Two additional dyes can be used, NED for OeduU2 and JOE for OeduH15 (Applied Biosystems). After PCR amplification, 1 µl PCR product was added to 10 µl formamide and 0.25 µl Rox500 size standard. Samples were denatured at 96°C for 5 minutes and then immediately cooled on ice. Multiplex loading allowed the analysis of 3 or 4 markers at the same time, by using markers labelled with different dyes or whose size range does not overlap. Electrophoresis parameters were set at injection for 15 s at 15 kv, running for 30 min at 15 kv and 60°C, with POP4 polymer. Data were analysed with GeneMapper® software version 3.7.

For *O. edulis*, 24 microsatellites have been published. *Oedu*.B0, *Oedu*.B11 and *Oedu*.C6 were identified by Naciri et al. (1995). *Oedu*G9 and *Oedu*T18 were described in Launey (1998) as suitable for use in segregating families. Oe1/10, Oe1/21, Oe1/47, Oe1/63, Oe1/64, Oe2/71, Oe3/37, Oe3/44 and Oe4/19 were developed by Morgan et al. (2000) and Morgan and Rogers (2001). *Oedu*.HA1, *Oedu*.HA7, *Oedu*.HA10, *Oedu*.HA11a and *Oedu*.HA21 were isolated by Sobolewska et al. (2001). Finally, *Oedu*J12, *Oedu*U2, *Oedu*H15, *Oedu*T5 and *Oedu*O9 were described by Launey et al. (2002). Primers pairs sequences, size range and PCR conditions as described in the literature are shown in Table 8.

Table 8. Microsatellites of *Ostrea edulis*. T_a: temperature of annealing; N_o: number, H_{obs} and H_{exp}: observed and expected heterozygosity, bp: base pair, R: reverse primer, F: forward primer. Numbers in brackets represent number of individuals that have been scored.

Locus	Microsatellite	Primers	$MgCl_2$ (mM)	T_a (°C)	N_o. of alleles	Size range (bp)	H_{obs}	H_{exp}
Naciri et al. (1995)								
Oedu.B0 *(72)*	$(GGT)_7GGGG(GC)_9$	R : ACC TTT ATA CAA ATC ATC CCT G F : ACA TGA CCC TTC ATT TGA ACC	1.0	55	5	90-102		0.660
Oedu.B11 *(73)*	$(TG)_9(AC)_{22}$	R : TTG CAT TCA CTA AAT ATA ACC F : GGA TTC TTA CTT ACC CAT C	1.0	51	48	155-217		0.970
Oedu.C6 *(51)*	$(TC)_{26}$	R : GTA TTA ATG AAA ATT GGA AGG F : GTC AAC TCC CAA CAC TG	1.0	53	17	131-175		0.910
Launey (1998)								
*Oedu*G9	$(AT)_4(GT)_{35}$	R : GAT CTT TGA TAA AGT ATT ACA GG F : GTT GAA CTC ACG ACC AAA TC	1.5	50		197-279		
*Oedu*T18	$(CT)_{21}$	R : CGC AAC TGA GTT ATG TAA AAT TG F : TTT AGT TCA TAT TTA GAA TAA ATC	1.5	50		-		
Morgan et al. (2000), Morgan and Rogers (2001)								
Oe1/10 *(67)*	$(CA)_5CCCACCACCCC(CA)_4$	F:ATCTGGGACTTTCGGGTTTC R:CATGTAACGGCTGCACAATC	1.0	62	3	197-201	0.269	0.281
Oe1/21 *(67)*	$(CA)_6(CG)_3CAAACAACACCACGC(CA)_2C(CA)_3$	F:AGGGAAAGGAAACGAGAATG R:ATGTCACGTGATGATGC	1.0	56	3	245-261	0.448	0.600
Oe1/47 *(67)*	$(GT)_{10}$	F:CAACAAGAAAACCACCATAC R:TGAATATGAGAAAGAAAACGAG	2.5	56	2	267-269	0.537	0.517
Oe1/63 *(67)*	$(GT)_9$	F:TAGATTCCTGGGCGATGG R:GGCTAAACTATCCCCTCACTG	1.0	60	7	96-112	0.731	0.772
Oe1/64 *(67)*	$(GT)_4GCGT(GC)_2(GT)_4AT(GT)_3$	F:TGTCGTTTCACAGTCATCGG R:CACCGTACCCTCCACGC	1.0	62	4	148-158	0.493	0.711
Oe2/71 *(67)*	$(CA)_4GA(TA)_3CATG(TA)_3C(GT)_8TT(GT)_3$	F:CCCATAGGCGCAGTGTTTAC R:AGAACCTTCCCTGCTCAATG	1.0	62	3	302-310	0.328	0.329
Oe3/37 *(67)*	$(GT)_{10}$	F:CCCGTTACATATAAAGGTATGTGTTG R:GGGCGCCTCCAAATAATG	2.5	56	3	217-221	0.313	0.389
Oe3/44 *(67)*	$(GT)_3GC(GT)_4AT(GT)_3ACGCAT(GT)_7$	F:CCAAATTAGCCACAAATCATCC R:CAGACTCATGGCGCAAAC	1.0	56	15	184-248	0.522	0.580
Oe4/19 *(36)*	$(GT)_5TT(GT)_7$	F:TCGGCACCGTACGTAATC R:TTTTGATAAAATTTCCGTTATTTG	2.5	50	1	160	0.000	0.000
Sobolewska et al. (2001)								
Oedu.HA1 *(29)*	$(GA)_{24}(AT)_8$	R : TTC CGT TGA TCC TCA TCT TG F : GAT CCC CAC GGG ATA CAA	1.5	57	17	144	0.966	0.908

Oedu.HA7 *(29)*	$(GA)_{28}$	R : TTC ACC CAG CCC TTC ATT TA F : TGT CTT TGA CAT ACG ATG ATG C	1.5	56	15	175	0.862	0.858
Oedu.HA10 *(29)*	$(AT)_5(GA)_{14}$	R : TGC ACT ACA TGT GCG TTA ATT F : GGT GAC GTC TCC ACA TGA GT	1.5	56	4	144	0.483	0.468
Oedu.HA21 *(29)*	$(GA)_{25}$	R : GAC CGT TTT TGC AGT CAG TG F : GCA CAT CCC TGG CAG TAT CT	1.5	68	13	184	0.857	0.891
Oedu.HA11a *(29)*	$(CA)_5TA(CA)_5(GA)_{17}$	R : ATC GTC GAT TTG GGC ATA AC F : CTG CGA TGC ACA AGA CTA A	1.5	67	14	200	1.000	0.904
Launey et al. (2002) 15 populations sampled								
*Oedu*J12	$(GT)_{14}$	R : TCG TCA CCT CCC TCT CAG AG F : GCT GTA TTT CCA TCA ATT CGA G	1.5	50		216-272		
*Oedu*U2	$(AC)_{21}(AG)_7$	R : GAA AGA AAT GGA GGC AAT AAC F : ACC AAT GAA CAC AGA TCA CC	1.5	50		146-206		
*Oedu*H15	$(\underline{ATCT})_4ATGT(\underline{ATCT})_2ATGT\underline{ATCT}ATAT\underline{ATCT}$ ATGT$(\underline{ATCT})_5A(TACC)_4AATTTTTCT(\underline{ATCT})_3$	R : TAA TGA TTT CGT TCG TTG AC F : TTT TGA CTC TGT GAT ATC GAC	1.5	50		165-225		
*Oedu*T5	$(CA)_{15}$	R : TAG TGA ATG GTC TTG CAT TCC F : CTT CGT TCT TGT ACG TAA GCG	1.5	55		106-166		
*Oedu*O9	$(GA)_{36}$	R : ACT TCA ATG TCT GTT CTA ATG G F : ATT CAA TTG ATT TTA GGT TGG	1.5	53		140-176		

Although all published methods indicated that microsatellite amplification had been optimised, nevertheless further optimisation was required. PCR conditions were modified by changing annealing temperature, primer concentrations, $MgCl_2$ concentration or type of cycles used. Several tests of multiplex PCR were carried out allowing *Oedu*J12, *Oedu*U2 and *Oedu*H15 to be amplified in a single PCR. Multiplex PCRs were also achieved for Oe1/63, Oe1/64, Oe1/10 and Oe2/71 and for *Oedu*.HA1 and *Oedu*.HA7. The development of multiplex PCR required adjustment of concentrations of each primer pair so that band intensities of all markers amplified in a single amplification reaction had roughly the same intensity. For details of PCR conditions, see Table 9 and Figure 12. Despite several attempts, amplification of microsatellite locus Oe1/21 could not be achieved. For *Oedu*.HA10 a single band of 110 bp was revealed, perhaps corresponding to the amplification of another locus and locus Oe4/19 was confirmed as monomorphic (Morgan et al., 2000). Therefore, these 3 microsatellite loci were not used.

Table 9. PCR conditions for amplification of microsatellite markers of *O. edulis*. PCR cycles notations refer to Figure 12. V_f: final volume of the PCR reaction, C_{DNA}: amount of DNA used as template, T_a: annealing temperature.

Locus	Dye	Special feature	V_f	C_{DNA}	T_a	$MgCl_2$	Primer	PCR cycles
*Oedu*J12	FAM	Multiplex	10 µl	200 ng	50°C	2.5 mM	0.3 pmol	Type 1
*Oedu*U2	NED						0.3 pmol	
*Oedu*H15	JOE						0.5 pmol	
*Oedu*T5	FAM	-	10 µl	100 ng	55°C	2.5 mM	0.5 pmol	Type 1
*Oedu*O9	FAM	-	10 µl	100 ng	55°C	2.5 mM	2.5 pmol	Type 1
*Oedu*G9	HEX	-	10 µl	200 ng	50°C	2.5 mM	0.6 pmol	Type 1
*Oedu*T18	FAM	-	10 µl	200 ng	50°C	2.5 mM	3 pmol	Type 1
Oedu.HA1	HEX	Multiplex	15 µl	300 ng	56°C	1.5 mM	5 pmol	Type 2
Oedu.HA7	FAM						1 pmol	
Oedu.HA11a	HEX	-	15 µl	200 ng	67-55°C	1 mM	10 pmol	Touchdown
Oedu.HA21	FAM	-	15 µl	200 ng	68°C	1.5 mM	10 pmol	Type 2
Oe1/47	FAM	-	15 µl	200 ng	56°C	2.5 mM	0.5 pmol	Type 2
Oe1/10	HEX	Multiplex	15 µl	300 ng	62°C	1 mM	4 pmol	Type 2
Oe1/63	FAM						2 pmol	
Oe1/64	HEX						3 pmol	
Oe2/71	FAM						4 pmol	
Oe3/37	HEX	-	15 µl	200 ng	56°C	2.5 mM	0.5 pmol	Type 2
Oe3/44	FAM	-	15 µl	200 ng	56°C	1 mM	2.5 pmol	Type 2
Oedu.B0	FAM	-	15 µl	200 ng	55°C	1 mM	0.6 pmol	Type 3
Oedu.B11	HEX	-	15 µl	200 ng	51°C	1 mM	0.8 pmol	Type 3
Oedu.C6	FAM	-	15 µl	200 ng	53°C	1 mM	0.6 pmol	Type 3

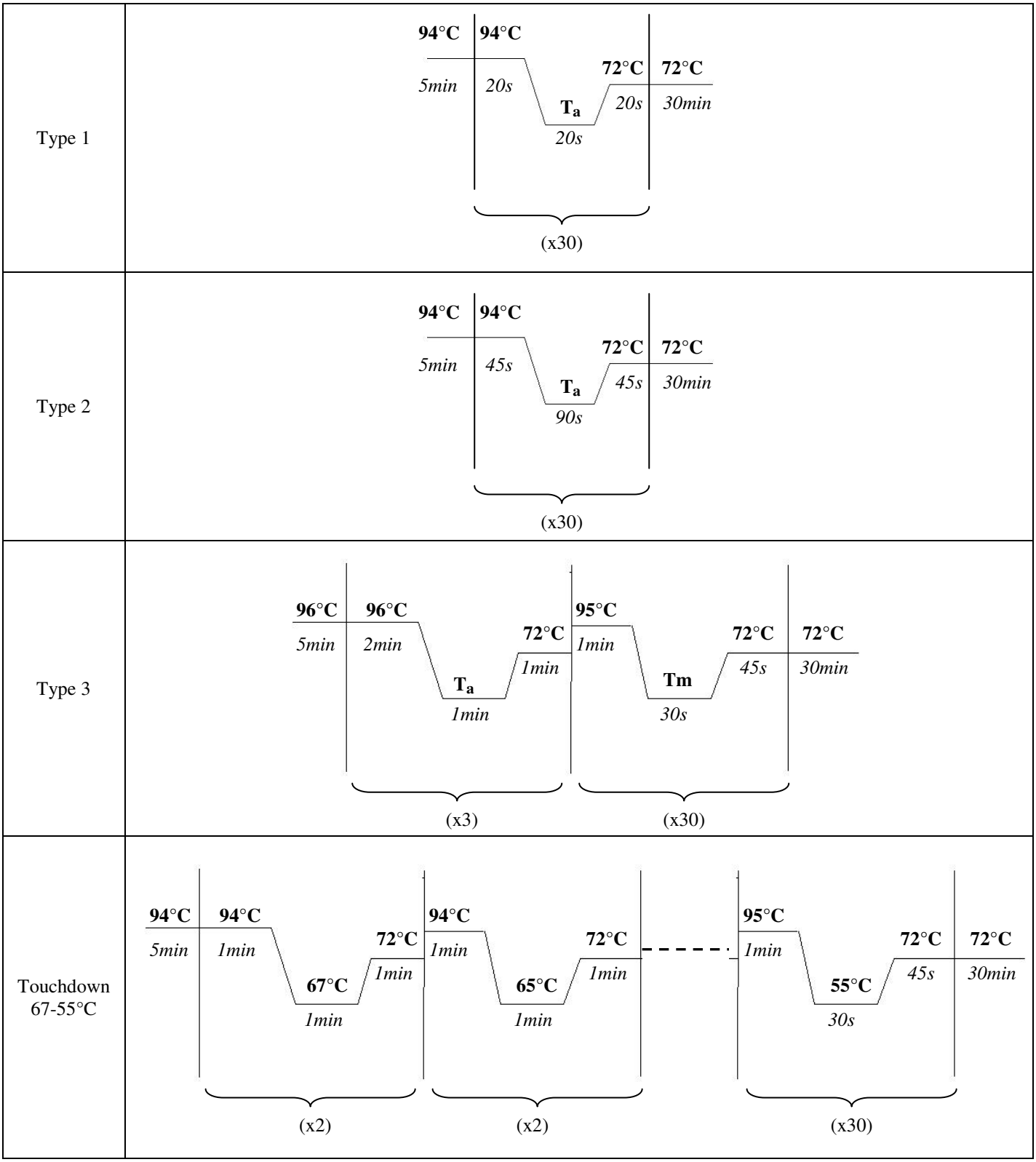


Figure 12. Thermocycler PCR programs used for the amplification of microsatellites of *O. edulis*.

For *M. edulis*, only 7 microsatellite markers have been published: *Mgμ1*, *Mgμ2*, *Mgμ3*, *Mgμ4*, *Mgμ5*, *Mgμ6*, *Mgμ7* (Presa et al., 2002). Primer pair sequences, size range and PCR conditions as described in the literature are shown in Table 10.

Table 10. Microsatellits of *Mytilus edulis* (Presa et al., 2002). See Table 8 for abbreviations.

Locus	Microsatellite	Primers	$MgCl_2$ (mM)	T_a (°C)	N_o. of alleles	Size range (bp)
Presa et al. (2002)						
Mgμ1 *(61)*	$(TG)_n$…$(AT)_n$	ATC AGA ATG GCA AAG AAA AA ACT ATG ATG GCT GAG AGG ATA	1.4	56	11	168-208
Mgμ2 *(58)*	$(CT)_n$	GGG ATC GTT CAA TAA GTT C AAA TTT TAC TGA ATA AAT AAA TCG	1.8	55	13	84-138
Mgμ3 *(82)*	$(TG)_n$	AAA CTA AAA ACT TCA TCT AAT CCC AAG CAA TCC AAA GTG AGA GG	1.8	60	5	143-151
Mgμ4 *(12)*	$(TG)_n$	CCT TAC TAT GCG TCG TTC AA TGA CCA ACA CTC CAA AAA TC	1.0	55	10	91-129
Mgμ5 *(16)*	$(TTTG)_n$	ACT TCT CCG GTA ACA TAA TA AGT CTT TCC CCT ATG ATG A	1.2	60	8	140-158
Mgμ6 *(36)*	$(AAT)_n$	GGG AAA GAC TGC CTA ACA AT CTC TTA CAT AGA AAA TGG TTC G	1.0	61	6	214-236
Mgμ7 *(43)*	(TA)n…$(TG)_n$	TAA GTT ATT GAT AGT TCG TTC C CAA AAC CAG TGT CAT ACA TAG	1.0	60	14	196-230

PCR conditions are different for the different microsatellites. Successful amplification (checked on 2% agarose gel stained with ethidium bromide) was achieved with an annealing temperature of 55°C (*Mgμ1*, *Mgμ2*, *Mgμ4* and *Mgμ7*) or 60°C (*Mgμ3*, *Mgμ5* and *Mgμ6*); an $MgCl_2$ concentration of 1mM (*Mgμ4*, *Mgμ5*, *Mgμ6* and *Mgμ7*), 1.4 mM (*Mgμ1*), 1.8mM (*Mgμ2* and *Mgμ3*); 5 pmol of each primer; 1 U *Taq* DNA polymerase (Promega). For all these markers, PCR was achieved in a final volume of 15 µl, containing 13 µl of reaction mixture and 2 µl of DNA (about 200 ng). Amplification protocol consisted of an initial denaturation at 95°C for 5 minutes, followed by 35 cycles of denaturing at 95°C for 1 minute, primer annealing at T_a (55°C or 60°C) for 40 seconds, extension at 72°C for 50 seconds followed by a final extension at 72°C for 15 minutes (Figure 13). Tests of amplification on the ABI sequencing machine was not as successful as promised in agarose gel. *Mgμ4*, *Mgμ5*, *Mgμ6* and *Mgμ7* revealed too many peaks (more than 2). Decreasing DNA concentration and using touchdown PCR have been tried. More optimisation needs to be done for these microsatellites for a reliable genotyping but due to the low number of microsatellites available and lack of time the optimisation process of *M. edulis* microsatellites was not developed further.

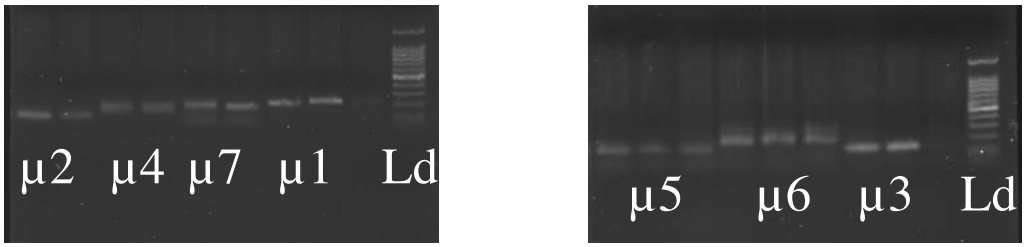


Figure 13. Successful amplification of *M. edulis* microsatellites checked on 2% agarose gels. Ld refers to molecular size standard (100bp DNA ladder, Promega): the strongest intensity band corresponds to 500bp.

Another locus has been amplified by PCR: mac-1, which corresponds to an intron-length polymorphism at the actin gene of *Mytilus* sp. (Daguin and Borsa, 1999). Specific PCR primers were used (forward 5'-CGT CTA GCG TAG TAC TTA TG-3' and reverse 5'-CGA AAA TTG TAG TCT AGT TTT GTG-3') as described by Daguin and Borsa (1999). PCR was carried out in 20 µl reaction mixture containing 300 ng template DNA, 0.5 µM of each primer, 1U *Taq* polymerase (Promega), 1X storage buffer A. Amplifications were achieved by running 30 cycles of 1.5 minutes at 94°C, 1 minute at T_a and 30s at 72°C following an initial denaturation of 5 minutes at 94°C and terminated by a final elongation step of 15 minutes at 72°C. The amplification of this locus is a good example of an optimisation process. We first used an $MgCl_2$ concentration of 2mM at an annealing temperature of 45°C but multi-banding was observed (Figure 14a). We then increased annealing temperature to 50°C to increase specificity and stringency of the PCR reaction: a single band was present (Figure 14b). At the same time, we tried another $MgCl_2$ concentration, 1.5 mM, and tested whether the presence of BSA (Bovine Serum Albumin) in the reaction mixture would result in a better yield of the reaction (stronger band intensity) (Figure 14c,d). Results showed that increasing T_a effectively suppressed multibanding, but $MgCl_2$ concentration or presence of BSA did not have an effect on the yield of the amplification (Figure 14).

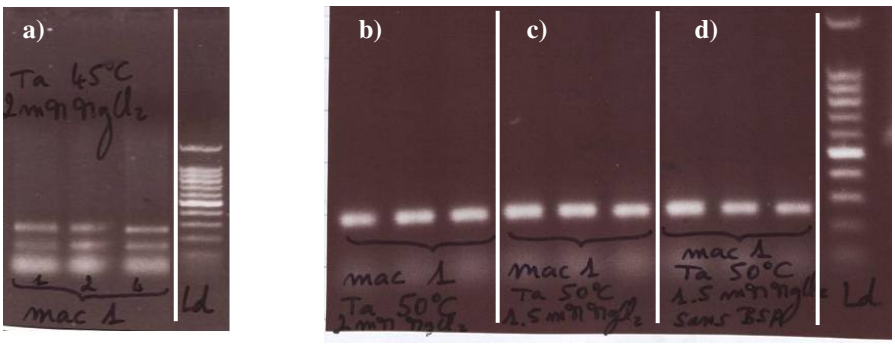


Figure 14. Optimisation of amplification of locus mac-1 in *M. edulis* checked on 2% agarose gels. a) T_a=45°C, $MgCl_2$ of 2mM; b) T_a=50°C, $MgCl_2$ of 2mM; c) T_a=50°C, $MgCl_2$ of 1.5mM, BSA; d) T_a=50°C, $MgCl_2$ of 1.5mM, no BSA. Ld refers to molecular size standard (100bp DNA ladder, Promega)

III.3. Optimisation of AFLPs in *O. edulis* and *M. edulis*

III.3.1. Optimisation of AFLPs methodology

The first steps of development of AFLPs were undertaken at the School of Ocean Sciences, University of Wales, Bangor. Scoring of AFLPs was done on a LI-COR® sequencing machine, on 6.5% denaturing acrylamide gels. After a pre-run of 25 minutes, runs lasted 3.5 hours with the following electrophoresis conditions: voltage 1500 V, power 40 W, current 40mA, temperature 45°C. The first protocol used was similar to the one of Vos et al. (1995), with restriction digest (4 hours at 37°C) and ligation (16 hours at 16°C) steps achieved separately. Optimisation was carried out using a single primer set with two selective nucleotides (Eco+CA and Mse+CG). Estimating the level of polymorphism was difficult because the banding pattern was complex and bands were often too close to each other for an easy interpretation of the gel. Therefore, new primers were ordered with three selective nucleotides, Eco+CAG and Mse+CGA.

All the further optimisation process was performed by using this primer pair (referred as A1). As expected, the complexity of the gel was far reduced by the addition of one selective nucleotide. Nevertheless very strong differences in band intensities sometimes made the analysis uncertain. The size range of the scorable bands was 50-350bp. To improve the ease of gel analysis and increase the size range of bands obtained, attempts were made to optimise the methodology.

Firstly, other DNA extraction methods were used: phenol-chloroform, CTAB chloroform without purifying DNA by the Wizard. However the success of amplification of AFLPs with these alternative DNA extraction methods was variable: sometimes it worked, sometimes not. The Promega kit extraction method was definitely the most reliable. Secondly, the number of cycles was increased during the selective PCR and a final elongation step was added at the end of the selective PCR. These 2 modifications did not significantly improve the quality of the gels obtained and artefactual bands were still present. Finally, following advice from Dr Tim Sharbel (Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany), both the restriction digest and the ligation were achieved in the same mix, incubating the mixture at 16°C for 16 hours. The aim was to correct for inconsistent DNA extraction problems between samples

(e.g. dilution) by extending the restriction digest step to ensure a complete digestion. Artefactual bands were reduced, the quality of reading improved considerably and band intensity was more homogeneous.

III.3.2. Assessment of repeatability of AFLP methodology

The repeatability of the technique was checked by comparing the band pattern of 4 replicates obtained independently (4 different DNA extractions, 4 different AFLPs amplifications, performed on different days) on the same 20 samples of *M. edulis*. The 4 replicates per sample consisted of DNA extracted from adductor muscle (2 replicates) or from gill (2 replicates). Band patterns obtained were identical; the AFLP technique is repeatable and no tissue-specific banding was observed (Figure 15). According to Matthes et al. (1998), different banding patterns could be the result of difference degree of methylation between tissues. However, although banding patterns were the same, bands intensities were lower for muscle samples and this confirmed the superiority gill-extracted DNA samples over muscle-extracted DNA.

III.3.3. Comparison of AFLPs obtained with 2 different sequencing machines

The development of AFLP markers, focusing mainly on the oyster *O. edulis*, was continued at Ifremer, La Tremblade, where 2 types of sequencing machines were available: a LICOR® DNA sequencer (6.5% acrylamide gel) and an ABI PRISM® 3100-Avant Genetic Analyzer (capillary). It seemed interesting to compare the banding patterns obtained on the same samples on those 2 machines. Therefore, independent AFLP amplifications were performed on the same 10 gill samples of *O. edulis*, whose DNA was extracted with the Promega kit. The method of visualisation of the banding pattern is different in the two systems, a band on slab gel for LICOR but a peak visible on a trace for ABI. Reassuringly, results obtained from the two systems were identical. The same bands/peaks were polymorphic at loci 72, 73, 74, 76, 88, 98, 131, 149, 151, 186, 217, 265 and 304 (named according to the size of the fragment in bp). Moreover, on the LICOR acrylamide gel, 2 replicates obtained independently were loaded and banding patterns in the two replicates were identical. This confirmed that the AFLP methodology is repeatable in *O. edulis* as well as in *M. edulis* (Figure 16).

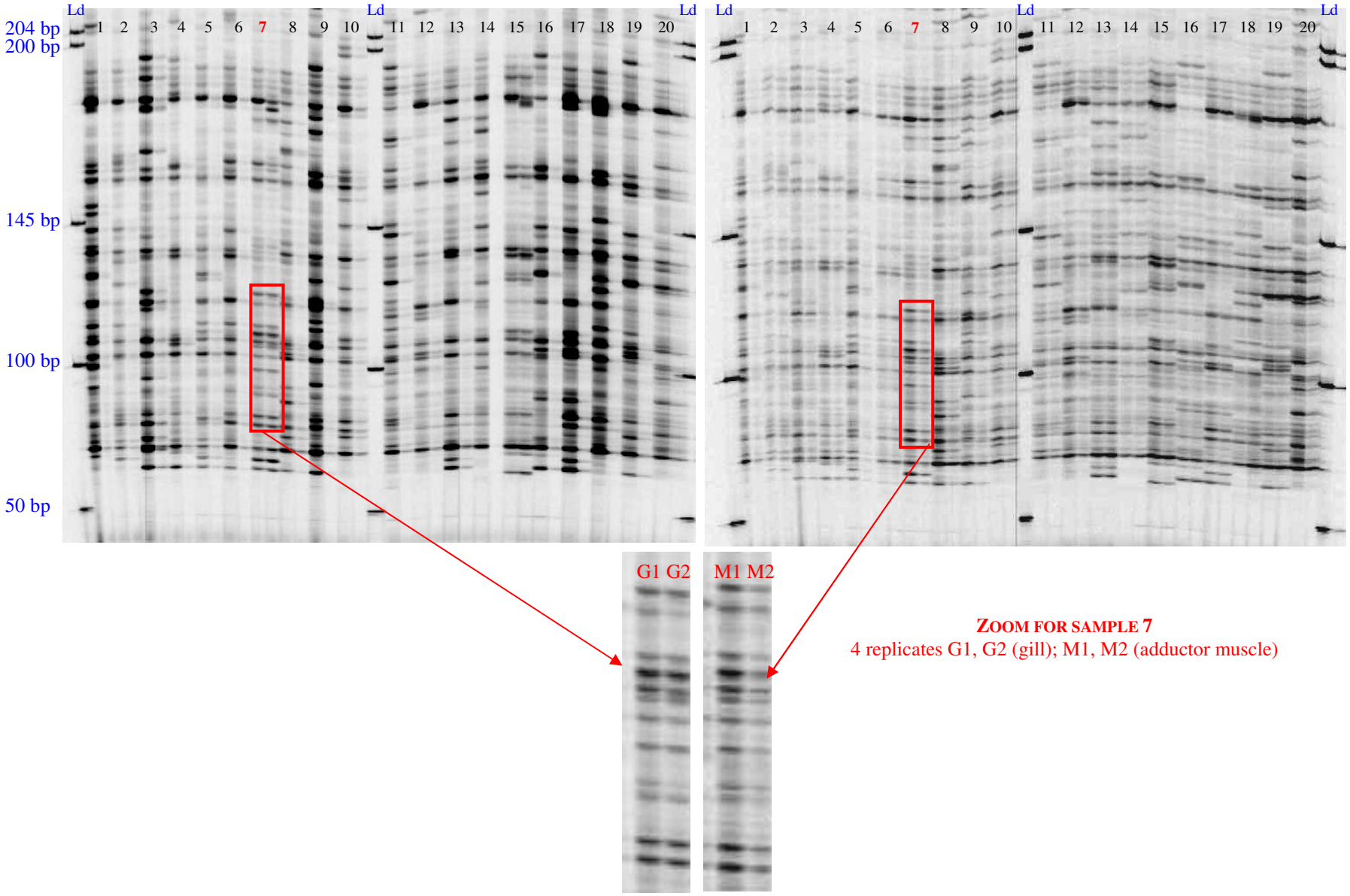


Figure 15. Assessement of the repeatability of the AFLP methodology. Four independent replicates were carried out on the same 20 samples of *M. edulis*, 2 for gill samples (left) and 2 for adductor muscle (right). Samples were loaded as followed: 1G1-1G2-2G1-2G2 (for the left picture, G1 and G2 being the 2 gill replicates); 1M1-1M2-2M1-2M2 (for the right picture, M1 and M2 being the 2 muscle replicates). Ld: molecular size standard.

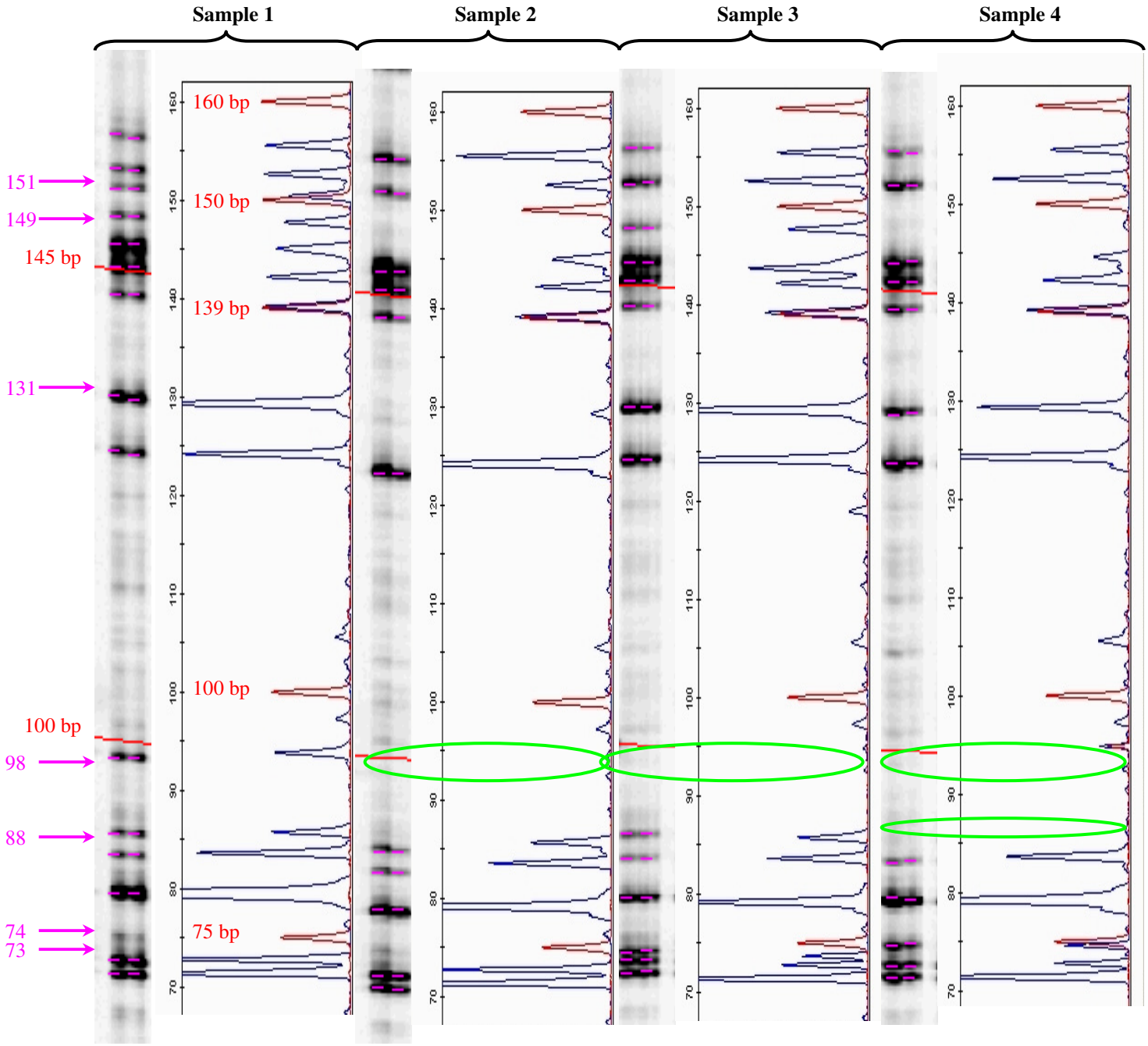


Figure 16. Comparison of AFLP patterns of 4 samples of *O. edulis*, visualised either on a LI-COR (6.5 % denaturing acrylamide gel) (left) or an ABI (capillary) (right) sequencing machine. Molecular weight standard are shown in red: 100 and 145 bp for LI-COR; 75, 100, 139, 150 and 160 bp for ABI. Pink arrows show polymorphic sites: 7 were identified, corresponding to fragments of size 73, 74, 88, 98, 131, 149 and 151 bp. Green ellipses show some examples where a fragment is absent, visualised by a lack of band or peak depending on the sequencing machine used, LI-COR or ABI respectively. For example, fragment 73 is present only in sample 3, fragment 74 in sample 4; fragment 88 is present in samples 1, 2 and 3; fragment 131 is present in samples 1, 3 and 4.

Although there was a good correspondence between the two systems, there was a difference of around 3 to 4 bp in size calling for the same loci visualised on the two different machines. For example, allele 98 of LI-COR corresponded to the one between 94 and 95 bp on ABI. Another interesting point concerns the background noise. On acrylamide gel, there is a low intensity smear which can be assimilated into the background noise (e.g., see Figure 16, between 100 and 125 bp). In contrast, background noise was almost absent on electrophoregrams obtained with the ABI machine. There were a few peaks of very low intensity that had to be discarded but generally speaking peaks were well defined and it was easier to analyse very close peaks (e.g. peaks of 71, 72, 73, 74 and 76 bp of the Figure 16). Therefore, all further AFLP analyses were performed on the ABI PRISM® 3100-Avant Genetic Analyzer due to the gain in quality and time. For the ABI system, there is no need to pour a gel and allow time for it to polymerise. More importantly, four different fluorochromes can be used on the ABI machine, one for the molecular size standard and three for the samples; this allows the multiplexing of three PCR reactions in one single run, in contrast to 2 for the LICOR machine where only 2 different wavelengths (700 or 800 nm) are available.

III.3.4. Final protocol used

Routine AFLP analysis of *M. edulis* and *O. edulis* DNA was performed using a modified version of Vos et al. (1995). Adaptors, preselective primers and non-labelled selective Mse+3 primers were purchased from Eurogentec. Selective EcoRI+3 primers were fluorescently 5' labelled with FAM, HEX or NED (Eurogentec and Applied Biosystems).

The digestion-ligation was carried out in a 50 µl reaction mixture containing 500 ng genomic DNA, 1X NEBuffer, 1X BSA, 1X T4 DNA ligase buffer, 5 U EcoRI, 5 U MseI, 200 U T4 DNA ligase (New England BioLabs), 50 pmol MseI adaptor, 5 pmol EcoRI adaptor. The reaction mixture was incubated at 16°C for 16 hours. Then 20 µl was diluted in 180 µl TE 0.1X.

Preselective amplification (PS-PCR) was performed in a final volume of 50 µl containing 5 µl diluted ligation-product, 1X storage buffer A, 1.5 mM $MgCl_2$, 0.1 mM dNTP, 25 pmol of EcoRI+1 and MseI+1 primers and 1 U *Taq* DNA polymerase (Promega). The PCR program for preselective amplification was 20 cycles of 94°C for 30 seconds, 56°C for 1 minute, and 72°C for

1 minute. Then 20 µl of PS-PCR product was diluted in 180 µl TE 0.1X. Selective amplification was performed in a final volume of 20 µl. Reaction mixture contained 5 µl diluted PS-PCR, 1X storage buffer A, 1.5 mM $MgCl_2$, 0.1 mM dNTP, 5 ng of EcoRI+3 and MseI+3 primers and 0.4 U *Taq* DNA polymerase (Promega). The PCR program consisted of the first 12 cycles in which the annealing temperature is reduced by 0.7°C each cycle, 94°C 30 seconds, 65°C 30 seconds, 72°C 1 minute. This is then followed by 23 cycles of denaturing at 94°C 30 seconds, annealing at 56°C 30 seconds and extension at 72°C 1 minute; finishing with a 30 minutes final extension at 72°C (See Table 11 for primer pairs sequences). Sixty primer pairs have been used (Table 12).

Table 11. Sequences of adaptors and primer pairs used for AFLPs methodology. The arrows indicate the restriction site of each enzyme. The part highlighted in yellow corresponds to the core sequence which helps stabilise the reaction; the one in pink corresponds to the sequence complementary to the enzyme restriction site; and the one in blue to the selective nucleotides added randomly to reduce complexity of banding pattern (one for PS-PCR and 3 for S-PCR).

	Eco RI enzyme	**MSe I enzyme**
Restriction site	G↓AATTC CTTAA↑G	T↓TAA AAT↑T
Adaptors	5′-CTCGTAGACTGCGTACC-3′ 3′-CATCTGACGCATGG TTAA-5′	5′-GACGATGAGTCCTGAG-3′ 3′-TACTCAGGACTC AT-5′
Primers	5′-GACTGCGTACC AATTC NNN-3′	5′-GATGAGTCCTGAG TAA NNN-3′

Table 12. Primer pairs used for scoring AFLPs, with their abbreviations.

	Eco+CAG FAM (blue)	**Eco+ACG HEX (green)**	**Eco+AAC FAM (blue)**	**Eco+ACA NED (yellow)**	**Eco+ACT NED (yellow)**
Mse+CGA	A1	B1	C1	D1	E1
Mse+CAA	A2	B2	C2	D2	E2
Mse+CTG	A3	B3	C3	D3	E3
Mse+CAT	A4	B4	C4	D4	E4
Mse+CTT	A5	B5	C5	D5	E5
Mse+ATC	A6	B6	C6	D6	E6
Mse+AGT	A7	B7	C7	D7	E7
Mse+CTC	A8	B8	C8	D8	E8
Mse+CTA	A9	B9	C9	D9	E9
Mse+CAC	A10	B10	C10	D10	E10
Mse+CAG	A11	B11	C11	D11	E11
Mse+CCT	A12	B12	C12	D12	E12

Electrophoresis and data collection was carried out on an ABI 3100-Avant (Applied Biosystems). The use of 3 different fluorochromes (FAM, HEX, NED) for the labelling of primer Eco+3 allowed the multiplexing of PCR products when running samples onto the sequencing machine. For example, analysis of pairs A1, B1 and E1 was simultaneous. After selective amplification, 3 µl PCR products (1 µl pair A1, 1 µl pair B1 and 1 µl pair E1) were added to 10 µl formamide and 0.25 µl Rox500 size standard (Applied Biosystems). Samples were denatured at 96°C for 5 minutes and then immediately cooled on ice to keep the DNA single-stranded. Electrophoresis parameters are set at injection for 15 s at 15 kv, running for 25 min at 15 kv and 60°C, with POP4 polymer. Data were analysed with GeneMapper® software version 3.7.

III.4. Feasability of developing SNPs (Single Nucleotide Polymorphism) in *O. edulis*

SNP (Single Nucleotide Polymorphism) markers are polymorphisms due to single nucleotide substitutions or single nucleotide insertions/deletions. In contrast with microsatellites, identification of SNPs is not based on length polymorphism, but implies to work on the sequence of DNA itself. The finding of SNP markers therefore relies on the amplification of a specific region of the DNA (codant or non codant) in several individuals by PCR and on the sequencing of the amplified fragment. After alignment of sequences obtained on different individuals, sequence comparisons allow the identification of SNPs by visualising nucleotide changes.

As only a few genes have yet been reported in *O. edulis*, attempts were made to amplify genes reported in another oyster species, the cupped oyster *C. gigas*. To increase the chances of amplification, touchdown PCR were performed for all primer pair tested, i.e. annealing temperature (T_a) was decreased by 2°C every two cycles.

Twenty primer pairs were first tested (Tanguy et al., 2004; Tanguy et al., 2005). PCR was performed in a final volume of 30µl, containing 200 ng template DNA, 0.02 mM dNTPs, 1X buffer, 5 pmol of each primer (forward and reverse) and 1 U of *Taq* polymerase (New England Biolabs). Two programmes were tested:

- 30 cycles of 30 s at 94°C, 1 minute at 58°C, 1 minute at 72°C, following an initial denaturation step of 2 minutes at 94°C and terminated by a final elongation of 10 minutes at 72°C
- Touchdown PCR 54-44°C similar to Touchdown 67-55°C of Figure 12, in which annealing temperature begins at 54°C and then decreases by 2°C every 2 cycles, to finish at 30 cycles with a T_a of 44°C.

Very few genes cross-amplified when using a fixed T_a of 58°C. However, when a touchdown PCR was performed, more genes amplified but the signal was quite weak (Figure 17a). To increase the intensity of the signal, a further 5 cycles at T_a 44°C were achieved in the touchdown PCR. Two different *Taq* enzymes were tested (New England Biolabs and Promega) and the signal was stronger using the Promega *Taq* (Figure 17b).

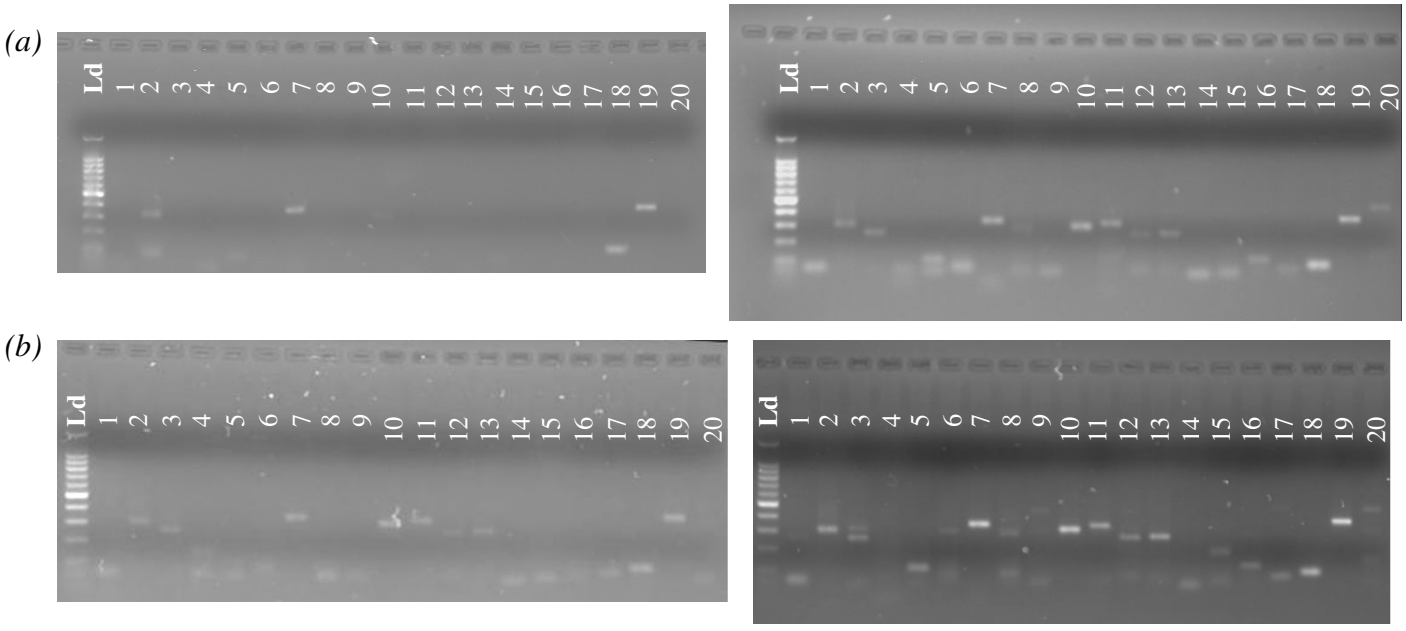


Figure 17. PCR amplification of genes of *C. gigas* in *O. edulis* under several conditions. a) Ta 58°C (left) and Touchdown PCR 54-44°C 30 cycles (right). b) Touchdown PCR 54-44°C 35 cycles with *Taq* polymerase New England Biolabs (left) or Promega (right). 1: apolipoprotein, 2: tripartite motif protein (TRIM), 3: actin, 4: lipoprotein receptor, 5: NADH oxidoreductase, 6: fucolectin, 7: senescente associated protein, 8: β 1-3 glucan, 9: proteasome, 10: Arha 2, 11: tetraspanin, 12: T-cell receptor, 13: heat shock factor 2, 14: beige protein, 15: RHO, 16: toll receptor, 17: glutamine synthetase, 18: glycogen phosphorylase, 19: nuclear corepressor, 20: calmodulin.

Of the 20 primer pairs tested derived from *C. gigas* (Tanguy et al., 2004 and 2005), most of them amplified in *O. edulis* but sometimes the signal was very weak (Figure 17b, right). Moreover, amplification for the *C. gigas* actin gene lead to a non-specific product as 2 bands can be visualised on the gel. Therefore, the 8 PCR products with the strongest signal were selected and used as templates for sequencing. These were the following *C. gigas* genes: tripartite motif protein (TRIM), senescente associated protein, Arha 2, tetraspanin, T-cell receptor, heat shock factor 2, beige protein, glycogen phosphorylase and nuclear corepressor.

Sequencing the PCR products required several steps. First, PCR products were purified with the Montage™ PCR Centrifugal Filter Device (Millipore): 380 µl distilled water and 20 µl PCR product was put into the sample reservoir and tubes holding the reservoir were centrifuged at 1000 x g for 15 minutes. Sample reservoirs were placed upright into a clean tube and 20 µl TE buffer was added to each tube. Tubes were then spun at 1000 x g for 2 minutes. The purified PCR products were used for the sequencing reaction itself. The sequencing reaction was carried out into a 6 µl final volume, containing 2 µl distilled water, 1.5 µl sequencing buffer, 0.5 µl BigDye® Terminator v3.1 (Applied Biosystems), 1 µl primer at 5 µM (primer F or R that served for the PCR) and 1 µl purified PCR product. The programme consisted of an initial denaturation of 3 minutes at 96°C followed by 35 cycles of 30 s at 96°C, 30 s at 55°C, 4 minutes at 60°C. This was performed using a 96-well plate. Sequencing reactions were then purified. In each sample, 64 µl 95% ethanol and 26 µl distilled water were added and the samples were spun at 4000 rpm for 30 minutes. The plate was inverted to remove ethanol. Then, 100 µl 70% ethanol was added, followed by a centrifugation of the plate at 4000 rpm for 10 minutes. The plate was then centrifuged upside down for 30 s to remove the ethanol. The elimination of ethanol is very important because residuals could induce artefactual peaks when running samples onto the sequencing machine (Sauvage, com.pers.). Finally samples were dried in a Speed Vac and re-suspended with 10 µl formamide.

Samples were loaded into an ABI PRISM® 3100-Avant Genetic Analyzer, using a 36 cm capillary array and POP 4 polymere. Of the 8 genes sequenced from one individual *O. edulis*, 7 exhibited very clear and clean sequences: TRIM, senescente associated protein, Arha 2 (unknown function), tetraspanin, T-cell receptor chain alpha c6.1A fusion protein, Heat Shock Factor 2 and

nuclear corepressor (Figure 18). Only one failed to produce a clear sequence, glygogen phosphorylase, probably because the length of the fragment amplified was too short, around 70 bp.

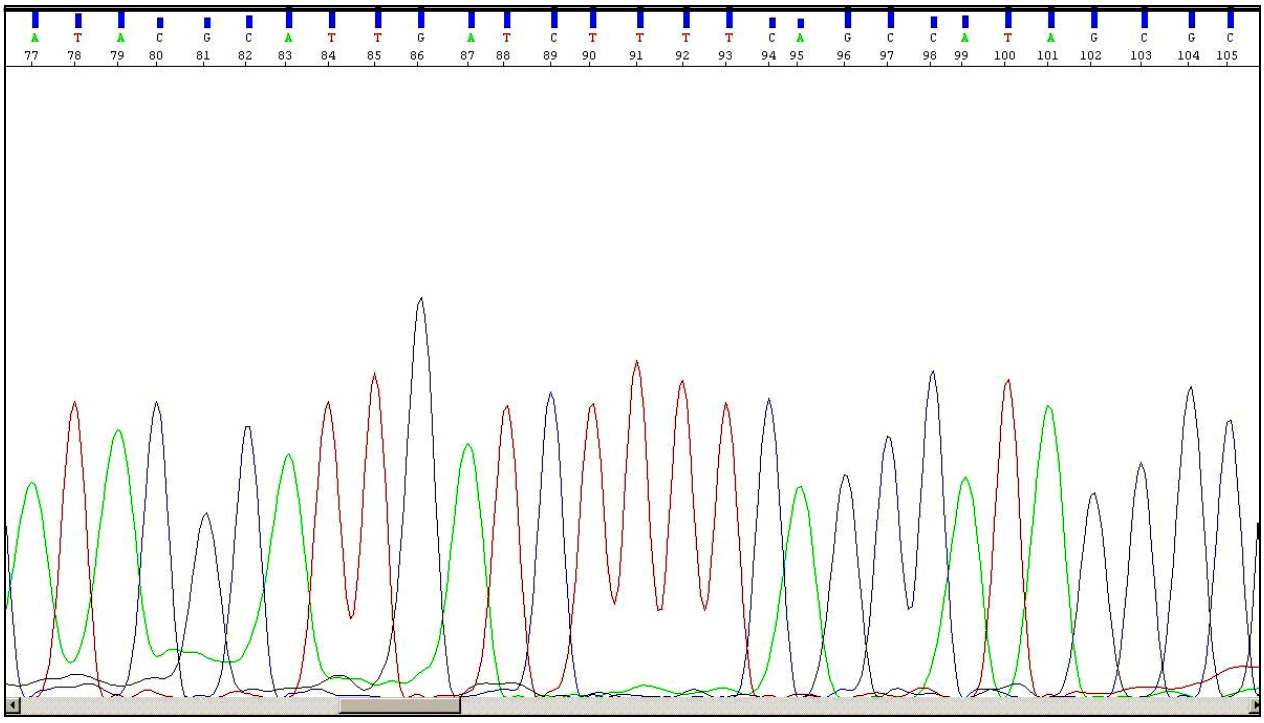


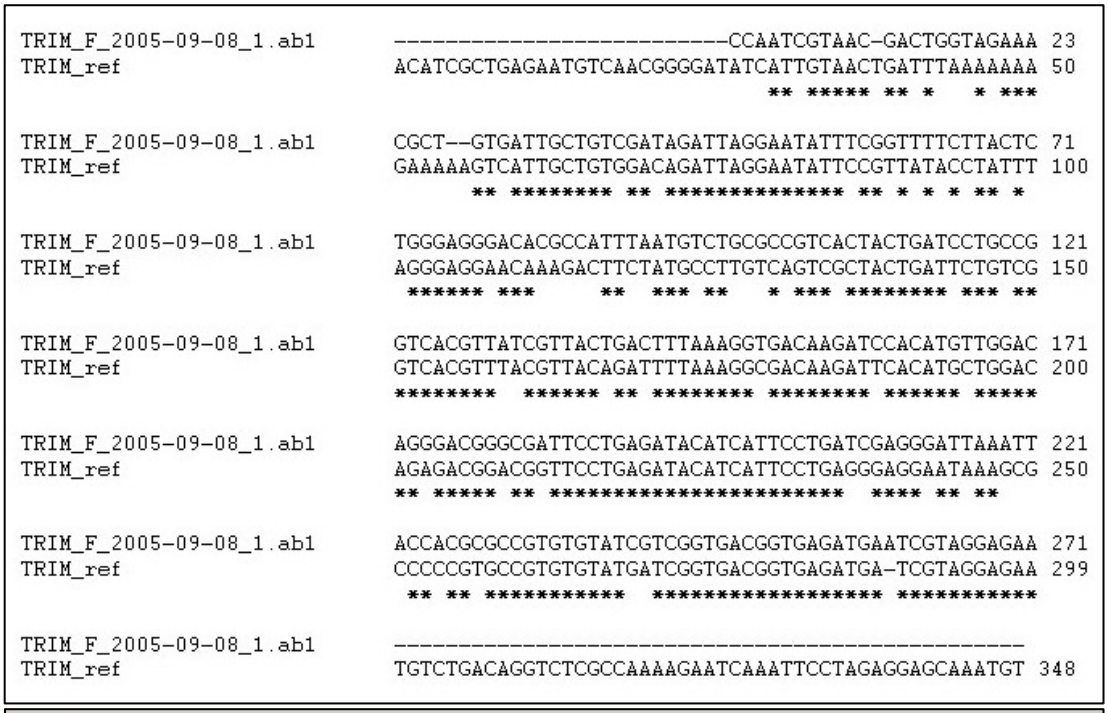
Figure 18. Partial sequence obtained in *O. edulis* after cross-amplification with primers developed in *C. gigas* for gene Arha 2. This sequence was obtained by using the forward primer during the sequencing reaction.

For the 7 sequences obtained, alignment with the reference sequence was performed with Clustalw (www.babel.infobiogen.fr). Three of them showed good (TRIM and Arha 2) or perfect (senescente associated protein) alignment with the reference sequence, confirming that the expected fragment of gene was amplified. Sequence alignments obtained for TRIM, Arha 2 and senescente associated protein are shown below. A star (*) corresponds to the same nucleotide between the reference sequence (bottom sequence) and the one amplified on *O. edulis* (top sequence). The absence of a star symbolises a nucleotide substitution between the two sequences or a deletion of one nucleotide (visualised by "-").

The sequence Senes_F obtained in the *O. edulis* sample therefore differs with the reference sequence at 6 positions, 2, 3, 5, 27, 47 328 bp.

```
Senes_F_2005-09-08_1.ab1      ----------------------------------------------ACG 3
Senes_ref                     ACTGAGCAGGATTACCGTTGCAACGACTGCAGTCATCAGTAGGGTAAAAC 50
                                                                             *

Senes_F_2005-09-08_1.ab1      TGACCTGTCTCACGACGGTCTAA-CCCAGCTCACGTTCCCTATTAGTGGG 52
Senes_ref                     TAACCTGTCTCACGACGGTCTAATCCCAGCTCACGTTCCCTATTGGTGGG 100
                              * ********************* ******************* *****

Senes_F_2005-09-08_1.ab1      TGAACAAGCCAACGCTTGGTGAATTCTGCTTCACAATGATAGGAAGAGCC 102
Senes_ref                     TGAACAAGCCAACGCTTGGTGAATTCTGCTTCACAATGATAGGAAGAGCC 150
                              **************************************************

Senes_F_2005-09-08_1.ab1      GACATCGAAGGATCAAAAAGCAACGTCGCTATGAACGCTTGGCTGCCACA 152
Senes_ref                     GACATCGAAGGATCAAAAAGCAACGTCGCTATGAACGCTTGGCTGCCACA 200
                              **************************************************

Senes_F_2005-09-08_1.ab1      AGCCAGTTATCCCTGTGGTAACTTTTCTGACACCTCTTGCTTAAAACTCC 202
Senes_ref                     AGCCAGTTATCCCTGTGGTAACTTTTCTGACACCTCTTGCTTAAAACTCC 250
                              **************************************************

Senes_F_2005-09-08_1.ab1      TAAAATCAAAAGGATCGATAGGCCCCGCTTTCACGGTCTGTATTCGTACT 252
Senes_ref                     TAAAATCAAAAGGATCGATAGGCCCCGCTTTCACGGTCTGTATTCGTACT 300
                              **************************************************

Senes_F_2005-09-08_1.ab1      GAAAATCAAAATCAAGTGAGCTTTTGCCCTTTTACTCTACGCGAGGTTTC 302
Senes_ref                     GAAAATCAAAATCAAGTGAGCTTTTGCCCTTTTACTCTACGCGAGGTTTC 350
                              **************************************************

Senes_F_2005-09-08_1.ab1      CGTCCCTCGCTGAGCTCAGA---------------------------- 322
Senes_ref                     CGTCC-TCGCTGAGCTCACCTTAGGACACCTGCGTTACCGTTTGACAGAT 399
                              ***** ************

Senes_F_2005-09-08_1.ab1      --------------------------------------------------
Senes_ref                     GTACCGCCCCAGTCAAACTCCCCGCCTGACACTGTCTTCAGAGCGGATCG 449


Senes_F_2005-09-08_1.ab1      ------------
Senes_ref                     CCCCCGGCCGC 460
```

More discrepancy was found with TRIM and Arha 2, with 65 and 33 nucleotide changes identified respectively.

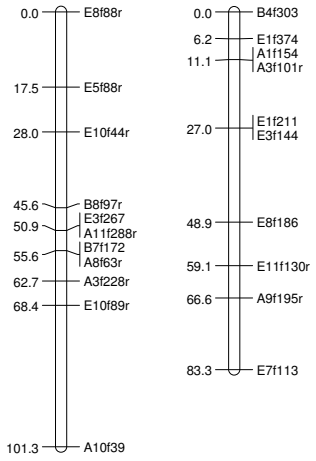


For the other 4 sequences, no good matches were found which suggests that a fragment other than the expected one was probably been amplified. No homology with sequences published in the database could be found with Blastn. However, this should not hamper the use of the cross-amplification approach to find SNPs in *O. edulis*. Indeed, SNPs being simply the search for nucleotide polymorphism, any sequence, whatever its function, is interesting as soon as polymorphism is exhibited.

In conclusion, due to the quality of the sequences obtained, it is feasible to use primers of genes developed in another species of oyster to identify more codominant markers in *O. edulis*. The next step would be to sequence the same fragment of gene in several individuals, or in the parents of controlled families to identify SNPs. This approach seems very promising. The only challenging step would be the genotyping itself as it requires complex and/or expensive systems, such as mass spectrometry, quantitative PCR. Here, only the feasibility of developing SNPs in *O. edulis* was investigated but we did not pursue this possibility further because of lack of time and lack of a SNP genotyping system in the laboratory.

CHAPTER 3

Genetic mapping in *Mytilus edulis*



I- INTRODUCTION

The blue mussel *Mytilus edulis* is a bivalve species of economical importance with production reaching 1 to 1.5 million tons a year (FAO, 2002). In addition, it is a model organism in marine biology and numerous biochemical, physiological and genetical genetics studies have been undertaken on this species (review in Gosling, 1992).

Genetic mapping consists of ordering molecular markers within the genome and estimating the relative distances between them. Construction of genetic linkage maps requires the use of highly polymorphic markers that are equally dispersed throughout the genome. Only 7 microsatellites are available so far in the literature for the blue mussel (Presa et al., 2002). Therefore, Amplified Fragment Length Polymorphism (AFLP) (Vos et al., 1995) markers were chosen as they require no preliminary knowledge of the genome, are highly reproducible (Jones et al., 1998) and can generate relatively quickly a high number of markers dispersed across the genome. The establishment of a genetic map relies on the relationship between recombination frequencies and genetic distances. The estimation of percentages of recombination between pairs of markers requires distinguishing between parental and recombinant gametes: a prerequisite of genetic mapping is to work with families with known parental genotypes. The experimental design chosen for this work on mussels was a full-sib family issued from an experimental biparental cross, with a double pseudo-test cross mapping strategy (Grattapaglia and Sederoff, 1994).

Despite their economical importance, no bivalve species have undergone any significant domestication, in contrast to agricultural animals and plants and bivalve production is based mostly on naturally collected seed. Nevertheless, some data on the genetic components of some traits of economical importance are available for the blue mussel. Significant heritability has been estimated for mussel traits such as growth and length (Mallet et al., 1986; Stromgren and Nielsen, 1989). Therefore, one value of establishing a genetic linkage map in *M. edulis* would be to contribute to a better understanding of the heritability of these and other traits. Linkage maps represent a framework which allows the localisation of QTL (Quantitative Trait Loci).

Identifying the number of QTLs and their localisation across the genome is a basic step in our understanding of the genetic basis of quantitative traits.

II- BIOLOGY OF *MYTILUS EDULIS*

II.1. Taxonomy

Organisms in the Phylum Mollusca generally consist of a soft, unsegmented body surrounded by an outer shell secreted by the mantle, and a strong muscular foot. In the class Bivalvia there is a bilaterally symmetrical pair of shell valves enclosing the body that lacks tentacles and a head. Large gills serve for respiration and food-collection and the foot serves for locomotion, digging or anchorage at different life stages. The order Filibranchia exhibit a pair of gills, each gill composed of an inner and outer demibranch which divides the pallial cavity into inhalant and exhalant chambers. Each demibranch comprises two lamellae (ascending and descending), held by interlamellar junctions. The organisation of junctions is not very complex (in contrast to other bivalves classes), represented by only ciliary tufts. The family Mytilidae is characterised by a solid shell with thin, elongated, equilateral valves and by the presence of byssus secreted by the foot in juveniles and adults. The posterior adductor muscle scar is large but the anterior adductor muscle scar is much reduced. The hinge lacks teeth but exhibits 3-12 small crenulations below the umbones.

The adult blue mussel *M. edulis* lives fixed to its support by the byssus. It can only move very slowly. Byssus gland secrets a liquid which polymerises with water and forms very resistant filaments that attach to the support on which mussel lives. The outer shell is approximately triangular in outline, smooth with a sculpturing of concentric lines but no radiating ribs, and its colour is purple, blue or sometimes brown (Beaumont et al., 1989). However, shell shape varies considerably with environmental conditions (Seed, 1968). Length varies, specimens usually ranging from 5-10 cm although some populations never attain more than 2-3 cm, and the largest specimens may reach 15-20 cm (Seed and Suchanek, 1992). Beaks are anterior, terminal, and rounded with straight umbones but not turned downward. The ligament runs from the beaks to the high point of the shell but is inconspicuous. The shell interior is pearl-white with a purple or

dark blue border. The edge of mantle is whitish yellow or brown (Beaumont et al., 1989). The flesh of the mussel is most often browny orange (female) or whitish (male) (Figure 19).

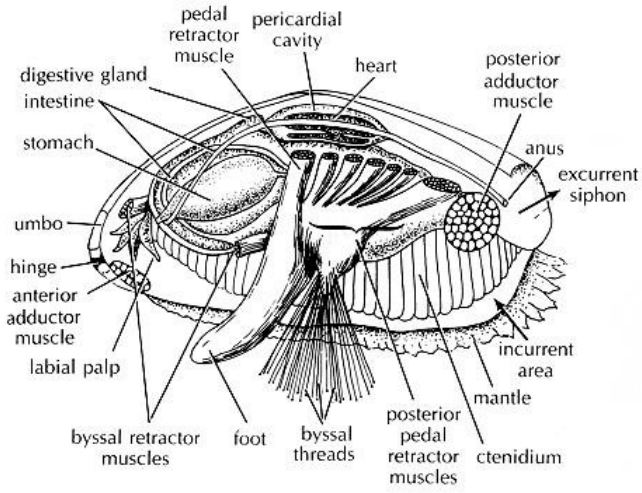


Figure 19. External and internal morphology of the blue mussel *Mytilus edulis*.
(http://www.glf.dfo-mpo.gc.ca/pe-pe/es-se/mussel-moule/images)
(http://www.manandmollusc.net)
(http://alpha2.bigelow.org/mitzi/lib_mid_an.html)

II.2. Geographic distribution and habitat

The genus *Mytilus* is widely distributed in northern and southern hemispheres, from temperate to boreal waters. Due to the difficulty of morphological species identification (which can be complicated by phenotypic plasticity in different environments), a combination of both morphological and genetic analyses is necessary for identifying *Mytilus* species with high confidence. *M. edulis* is widely distributed in the northern hemisphere, from the White Sea to the Atlantic coast of southern France. Electrophoretic analyses showed evidence of the absence of *M. edulis* on Northern Pacific coasts, but its presence in the western Atlantic with a restricted distribution on the east coast of North America, extending from the Canadian maritimes to Cape

Hatteras in North Carolina. This species also occurs in Iceland and probably in South America (Gosling, 1992). *M. galloprovincialis* has a more southerly distribution in Europe and can mostly be found in the Mediterranean Sea, the Black Sea, the Atlantic coasts of France, Spain and Portugal. This species has also been identified in southern California, Japan, Hong Kong, East China, south Africa, Australia, Tasmania and New Zealand (Gosling, 1992). *M. trossolus* is distributed in the northern hemisphere, in the Baltic Sea, the west coast of North America and the east coast in the region of the Canadian Maritimes (Gosling, 1992). *M. californianus* extends along the Pacific coast of North America, from Alaska to Mexico (Gosling, 1992) (Figure 20).

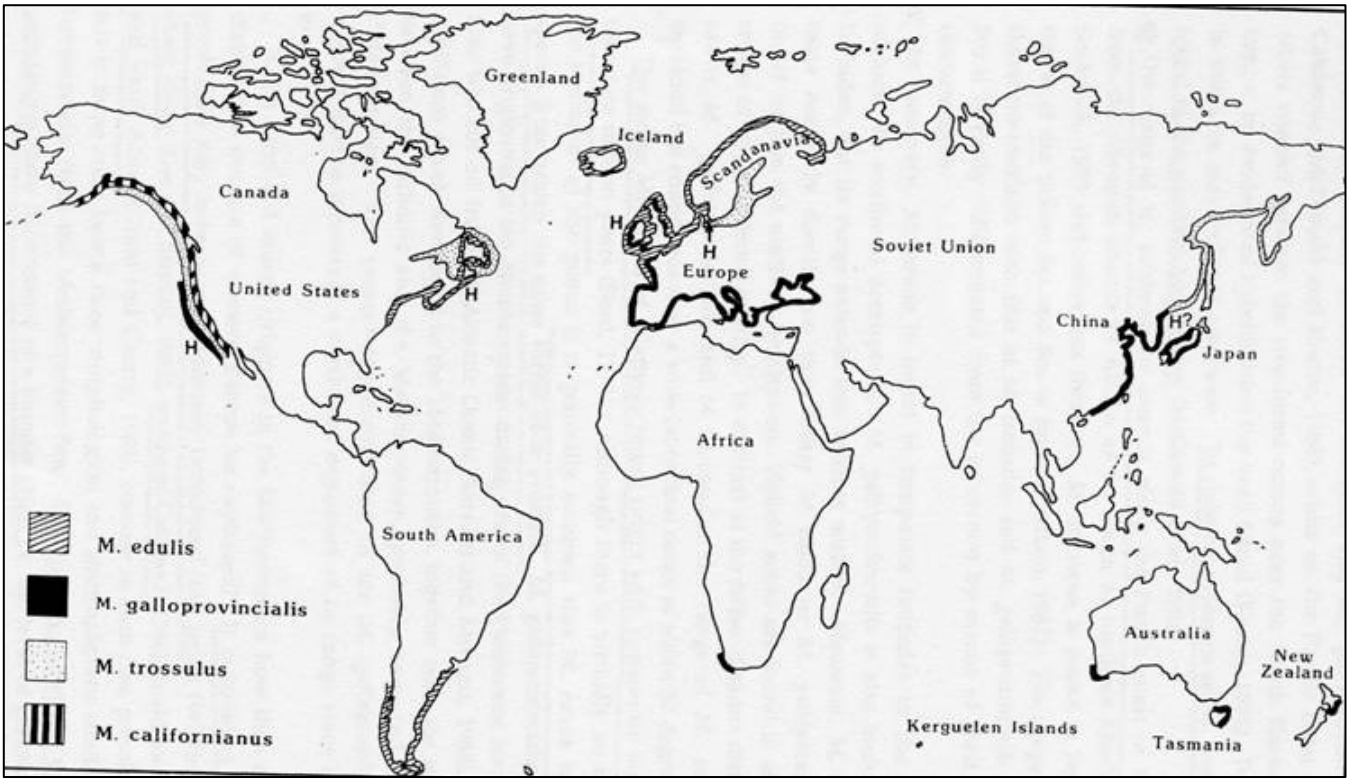


Figure 20. Geographic distribution of the genus *Mytilus* (Gosling, 1992).

M. edulis occurs from the high intertidal to the shallow subtidal attached by fibrous byssus threads to suitable substrata. Although sometimes abundant in the subtidal, *M. edulis* is primarily an intertidal species (Seed and Suchanek, 1992). It can withstand extreme wave exposure. The upper limit of *M. edulis* populations on rocky shores is determined by its tolerance to temperature and desiccation: sudden mass mortalities at the upper limit of intertidal mussel beds are often associated with prolonged periods of unusually high temperatures and desiccation

stress (Suchanek, 1985). *M. edulis* is relatively tolerant of extreme cold and freezing; it can survive occasional, sharp frost events, but may succumb to consistent very low temperatures over a few days (Williams, 1970; Bourget, 1983).

M. edulis can be found on the rocky shores of open coasts attached to the rock surface and in crevices, and on rocks and piers in sheltered harbours and estuaries. It is a gregarious species, and at high densities forms dense beds of one or more (up to 5 or 6) layers, with individuals bound together by byssus threads. Young mussels colonize spaces within the bed increasing the spatial complexity, and the bed provides support for rich epifaunal and infaunal invertebrate assemblages (Suchanek, 1978; Tsuchiya and Nishihira, 1986) (Figure 21).

Figure 21. Mussel bed showing an assemblage of associated organisms.
(http://www.marlin.ac.uk/species/Mytilusedulis.htm)

II.3. Population dynamics

Several factors can contribute to mortality and the dynamics of *M. edulis* populations. These include temperature, desiccation (Suchanek, 1978), storms and wave action (Harger and Landenberger, 1971), siltation (Ceccherelli and Rossi, 1984) and biodeposits (Tsuchiya, 1980), intra- and interspecific competition (Lewis, 1977; Richardson and Seed, 1990), and predation, which is the single most important source of mortality. Many predators target specific sizes of mussels and, therefore, influence population size structure. The main predators are neogastropods such as *Nucella lapillus*, starfish such as *Asterias rubens*, the sea urchin *Strongylocentrotus*

droebachiensis, crabs such as *Carcinus maenas* and *Cancer pagurus*, fish such as plaice, flounder and dab, and birds such as oystercatcher. The polychaete *Polydora ciliata* may burrow into the shell of *M. edulis*, which weakens the shell leaving individuals more susceptible to predation by birds and shore crabs (Seed and Suchanek, 1992). Several infectious diseases and parasites affecting mussels (including *M. edulis*) have been reported. However, most parasites do not cause significant mortality but several species of parasite found in mussels can also infect and cause mortality in other shellfish, e.g *Marteilia refringens* infecting the flat oyster *Ostrea edulis* (Bower, 1992; Berthe et al., 2004).

II.4. Feeding

Mussels are suspension feeders, filtering plankton and other small particles from the water column (Figure 22).

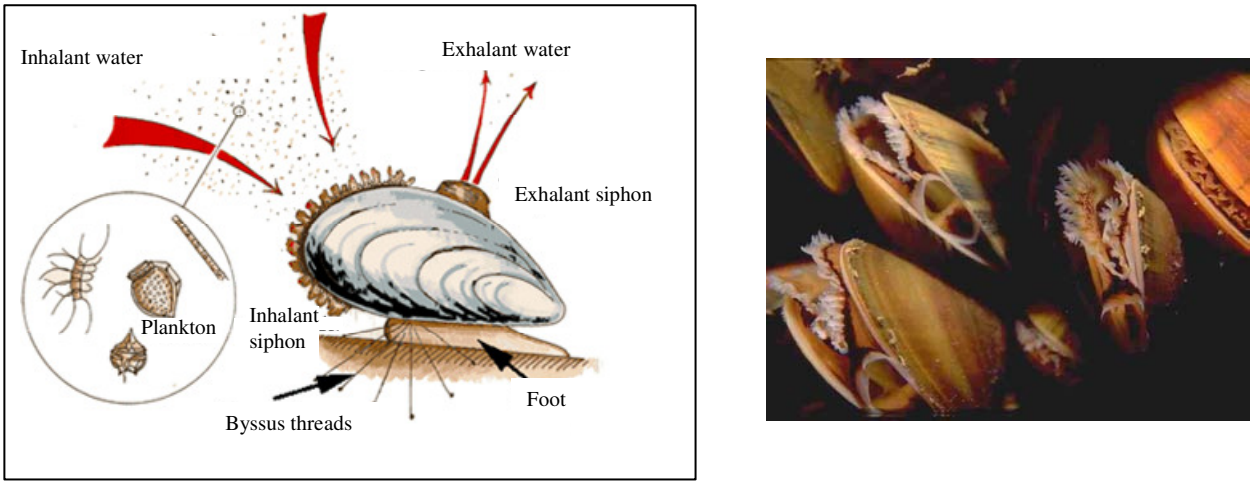


Figure 22. Feeding mode of the blue mussel *M. edulis*, shown in a diagram and a picture.
(http://www.glf.dfo-mpo.gc.ca/pe-pe/es-se/mussel-moule/mussel-moule-f.html)

II.5. Methods of cultivation

The methods used in the cultivation of mussels are either seabed or off-bottom cultivation. In seabed cultivation, seed mussels are transferred into culture plots for growth and fattening. Off-bottom cultivation includes pole cultivation (for example the bouchot style common in France) or the methods utilising ropes or lines suspended from the sea surface (Figure 23).

SEABED CULTIVATION

OFF-BOTTOM CULTIVATION

- Pole cultivation (mostly bouchot type)

- Ropes or lines suspended from the sea

Fixed suspended

Floating suspended (rafts, longlines)

Figure 23. Methods of cultivation of the blue mussel *M. edulis*.

II.6. Reproductive cycle

The sexes are almost always separate because mussels are dioecious. Gametogenesis and spawning varies with geographic location, e.g. southern populations often spawn before more northern populations. Reproductive output is influenced by temperature, food availability and tidal exposure and can therefore vary from year to year (Seed and Suchanek, 1992). An individual female can produce 7-8 million eggs, while larger individuals may produce as many as 40 million eggs (Thompson, 1979). Gametes are released in the water column where the fertilisation takes place. Fertilisation can occur successfully between 5-22°C and at salinities of 15-40psu (Bayne, 1965). Fertilised eggs are 60-90µm in diameter. The fertilised eggs differentiate into free-swimming larvae in 4-5 hours, into the trochophore larvae (unshelled, 24-48 hours after fertilisation) and then into the D-shape larvae (Jablonski and Lutz, 1980).

Dispersal is dependant on the duration of planktonic life. Maintenance of their position in the water column by active swimming ensures that larvae can be potentially dispersed over great distances by currents. In addition, post-larvae can become bysso-pelagic up to 2-2.5 mm in size, which may take around 2 months to achieve, during which time they may be transported significant distances by currents (Lane et al., 1985; Lutz and Kennish, 1992).

Larvae are free-swimming for 3 to 4 weeks until they seek to settle. At this stage, pediveliger larvae exhibit an eyespot, a foot, have lost their D-shape and have a pronounced umbo. Pediveligers typically settle at about 260 µm (McGrath et al., 1988) but can delay settlement for several weeks in absence of suitable substrata or optimal conditions (Bayne, 1971). Pediveligers test the substrata using their sensory foot. During metamorphosis, the mussel larvae extend the foot and are able to secrete a range of byssus threads for attachment or drifting (Lane et al., 1985), the velum breaks down and gills replace the velum for feeding. Primary settlement tends to occur on filamentous substrata, such as bryozoans, hydroids, filamentous algae (Eyster and Pechenik, 1987; King et al., 1990) or the byssus threads of previously settled adults. Post-larvae may remain on their primary attachment until 1-2mm in size (sometimes larger), and then will move by byssus drifting on adult beds for permament attachment (Lane et al., 1985). Primary

settlement may allow the pediveligers to avoid competition for food with adults or being inhaled by suspension feeding adults. Newly settled mussels are termed 'spat' or juvenile (Figure 24).

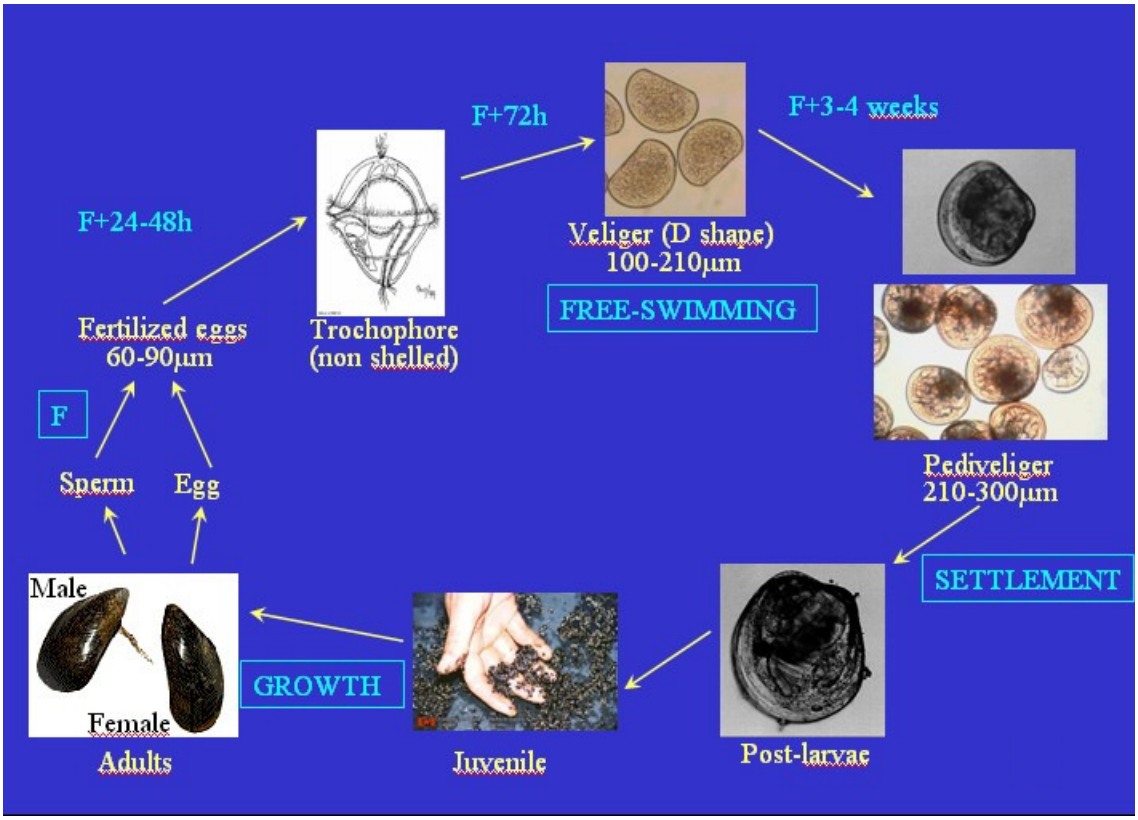


Figure 24. Reproduction cycle of the blue mussel *M. edulis*. Note the presence of gills in post-larvae.

II.7. Cytological features

Chromosomes of *M. edulis* have been studied at the morphological level by Giemsa-staining of the karyotype to estimate the number of chromosomes, as well as their visible size or shape. The chromosomes are characterised as metacentric, submetacentric, subtelocentric, telocentric and acrocentric, depending on the position of the centromere. All studies agree on the diploid number of chromosomes of *M. edulis*: 2n=28. Most studies also agree on the presence of 6 pairs of metacentric chromosomes. For the other 8 pairs of chromosomes, their classification varies according to the authors between submetacentric, subtelocentric and submetacentric/subtelocentric (Table 13). This variation is possibly dependant on the population studied but more likely on the degree of condensation of chromosomes (Thiriot-Quiévreux, 2002), and on artefactual difficulties. A karyotype of *M. edulis* is presented Figure25.

More recently, banding techniques have been employed on *M. edulis* for a better characterisation of chromosomal morphology and to obtain of information about heterochromatin patterns: silver-staining of NORs (Nucleolar Organiser Regions), C-banding and fluorochrome staining (Table 13). Nucleolar organizer regions (NORs) are loops of DNA that encode ribosomal RNA and are considered important in the synthesis of proteins. NORs can be selectively stained by a silver colloid technique and can be visualized as black dots under the transmission microscope. These visualized structures are commonly termed AgNOR dots. The number of AgNORs rises with the increasing proliferative activity of cells. NOR silver staining allows location of the ribosomal genes, because they are associated with specific acidic proteins reflecting their transcriptional activity. The reaction is based on silver binding to acidic nucleolar proteins (Moreno et al., 1985). C-banding highlights the position of constitutive heterochromatin (which remains densely-coiled throughout interphase) and relies on the differential reannealing of portions of DNA after phases of fixation, dehydration, denaturation and renaturation.

Finally, some studies have estimated the genome size of *M. edulis*. Hinegardner (1974) estimated the diploid DNA content to be 3.2 pg by Bulk Fluometric Assay on sperm cells, for a total of 9 individual assays (3 times for each 3 concentrations of sperm). Rodríguez-Juíz et al. (1996) estimated the genome size at 3.42 pg by flow cytometry, which corresponds to the mean obtained on 20 samples. These values are quite high in comparison with other bivalve species such as oysters.

Table 13. Morphological and banding features of chromosomes of *M. edulis* (Thiriot-Quievreux, 2002).
Ag-NORs: silver-stained Nucleolar Organizer Regions; m: metacentric, sm: submetacentric, st: subtelocentric
CMA3: chromycin A3

2n	Karyotype (Giemsa-staining)	Ag-NORs (chromosome location)	C-banding	Fluorochrome staining	References
28	4 m, 2 m/sm, 4 sm/st, 4 st	-	-	-	Thiriot-Quiéveux and Ayraud (1982)
28	6 m, 4 sm, 4 st	-	-	-	Thiriot-Quievreux (1984a)
28	6 m, 6 sm, 2 st	1-4 (8 and 11)	-	-	Cornet (1993)
28	6 m, 8 sm/st	1-4 (8 and 11)			Insua et al. (1994)
28	6 m, 8 sm/st	2-4 (6 and 7)	+	CMA3: 11 blocks	Martinez-Lage et al. (1995)
28	6 m, 8 sm/st	+	+	+	Martinez-Lage et al. (1996)

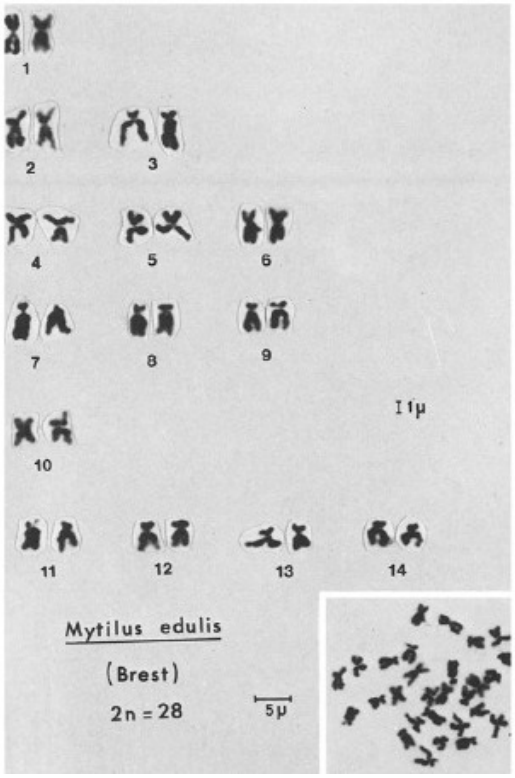


Figure 25. Karyotype established in a population of *M. edulis* coming from Brest (Brittany, France) (Thiriot-Quiévreux and Ayraud, 1982)

III- MATERIAL AND METHODS

III.1. Production of segregating families

Single pair matings of the blue mussel *M. edulis* were achieved at School of Ocean Sciences (University of Wales, Bangor, UK) with the collaboration of Dr Richard Braithwaite. Spawning was induced by injecting 1 ml KCl 0.5 M into the mantle cavity and holding mussels in air for 2 hours. Then, each injected mussel was placed individually into a jar, in which 1 µm-filtered sea water was added. Release of gametes was usually observed after several hours (up to 10 hours!!). Fertilisation occurred in 1 l cylinder by mixing eggs from one female and sperm from one male, with a ratio of approximately 10 spermatozoa per egg. This ratio was checked by observing 3 drops of 10 µl of the mixture under the light microscope. After the observation of the extrusion of the first polar body (5-10 minutes after fertilisation), around 100,000 fertilised eggs were transferred into a Pyrex dish (20 cm of diameter). Up to 3 replicates of fertilised eggs per single pair mating were kept when possible. After 72 hours, the D-shaped larvae were transferred into 2 l plastic beakers for the larval rearing process which lasted for 3 to 4 weeks before larvae

were approaching metamorphosis. During the larval stage, larvae were fed three times a week with a mixture of 2 phytoplankton species, 75% *Pavlova lutheri* (Prymnesiophyceae) and 25% *Rhinomonas reticulata* (Cryptophyceae), for a final concentration of 50 cells per µl (Figure 26). Sea water was changed three times a week: sea water was filtered on 1 µl and UV-light treated. It is important to note that all material (jars, beakers, test tubes, plungers, Pyrex dishes) were sterilised the day before by placing them into a water tank containing Chloros (a commercial solution of sodium hypochlorite) for 1 hour and by then rinsing each item in a tank containing clean fresh water during at least 2-3 hours. All these precautions were employed to avoid bacterial infection and obtain "good-quality" sea water.

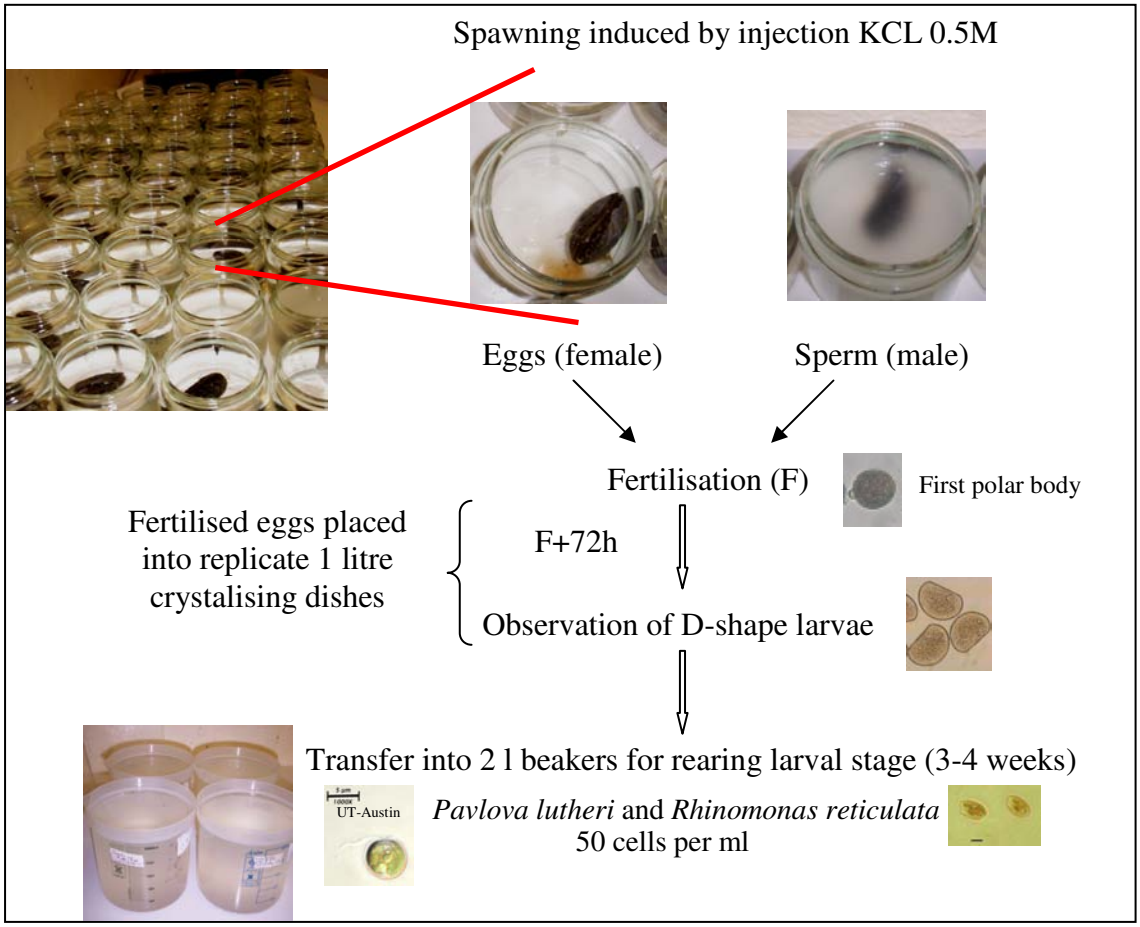


Figure 26. Experimental protocol for the production of segregating families of *M. edulis* by single-pair matings. From the induction of spawning to the larval rearing.

Several initial attempts were unsuccessful: the 14 single-pair mating families produced on the dates 11/05/2004, 21/05/2004 and 28/05/2004 were discarded at 3 days post-fertilisation because larvae obtained were either dead or abnormal. Thirteen further families were produced (named Family 1 to Family 13) on 4 different dates: 08/06/2004, 15/06/2004, 21/07/2004 and 23/07/2004. Ten of these 13 families were maintained through the larval rearing stages; families 1, 3, 4, 5, 6, 7, 10, 11, 12 and 13 (Table 14).

Table 14. Data on attempts to obtain segregating families in *M. edulis* to rear through the larval stage.

Date spawning	N adults	N families	N eggs.l^{-1}	Male	N replicates per family	Family kept	Reasons
11/05/2004	50	5	♀a: 4.10^6	♂a	3 (100.000 eggs)	No	No D larvae at day 3
			♀b: 6.10^5	♂b	3 (100.000 eggs)	No	No D larvae at day 3
			♀c: 1.10^5	♂c	2 (50.000 eggs)	No	No D larvae at day 3
			♀d: 4.10^5	♂d	3 (100.000 eggs)	No	No D larvae at day 3
			♀e: 5.10^5	♂e	3 (100.000 eggs)	No	Bacterial infection at day 6
21/05/2004	50	5	♀f: 2.10^6	♂f	3 (100.000 eggs)	No	Dead larvae at day 3
				♂g	3 (100.000 eggs)	No	Abnormal larvae at day 3
				♂h	3 (100.000 eggs)	No	Dead larvae at day 3
				♂i	3 (100.000 eggs)	No	Dead larvae at day 3
			♀g: 1.10^5	♂f	1 (100.000 eggs)	No	Dead or abnormal larvae at day 3
28/05/2004	37	4	♀h: 6.10^5	♂j	3 (100.000 eggs)	No	Abnormal larvae at day 3
			♀i: 4.10^5		3 (100.000 eggs)	No	Abnormal larvae at day 3
			♀j: 1.10^5		1 (100.000 eggs)	No	Abnormal larvae at day 3
			♀k: $2.7.10^5$		2 (100.000 eggs)	No	Abnormal larvae at day 3
08/06/2004	64	3	♀1: 3.10^5	♂1	3 (100.000 eggs)	Yes	10% D larvae at day3 (**Family 1**)
			♀2: 3.10^5	♂2	3 (100.000 eggs)	No	Too few D-shaped larvae (**Family 2**)
			♀3: 7.10^5	♂3	3 (100.000 eggs)	Yes	25-30% D larvae at day 3 (**Family 3**)
15/06/2004	64	3	♀4: 5.10^5	♂4	3 (100.000 eggs)	Yes	90-95% D larvae at day 3 (**Family 4**)
			♀5: 7.10^5	♂5	3 (100.000 eggs)	Yes	90-95% D larvae at day 3 (**Family 5**)
				♂6	3 (100.000 eggs)	Yes	90-95% D larvae at day 3 (**Family 6**)
21/07/2004	60	3	♀8: $1.2.10^6$	♂7	3 (100.000 eggs)	Yes	Few D larvae at day 3 (**Family 7**)
				♂8	3 (100.000 eggs)	No	Abnormal larvae at day 3 (**Family 8**)
			♀9: $2.5.10^5$	♂9	2 (100.000 eggs)	No	Abnormal larvae at day 3 (**Family 9**)
23/07/2004	56	4	♀10: 7.10^5	♂10	2 (100.000 eggs)	Yes	70% D larvae at day 3 (**Family 10**)
				♂11	2 (100.000 eggs)	Yes	70% D larvae at day 3 (**Family 11**)
			♀11: 5.10^5	♂12	2 (100.000 eggs)	Yes	90% D larvae at day 3 (**Famille 12**)
			♀12: 1.10^6	♂13	2 (100.000 eggs)	Yes	90% D larvae at day 3 (**Famille 13**)

All the larvae of Family 7 died at the end of the larval rearing, at around 1 month-old. Otherwise, "ready-to-metamorphose" larvae (evidence of eyed larvae and some pediveligers) from the other 9 families were transferred into post-settlement systems.

Two different systems were used. First, larvae from families 1, 3, 4, 5 and 6 were transferred into a downwelling system, on 85 µm-mesh sieves held in a big tank. Seawater flowed into the top of each sieve. Food was added manually every day by cutting the water flow and pouring phytoplankton (*P. lutheri*, *R. reticulata*) into each sieve (Figure 27a). As the seawater was not filtered, after a few days, sieves became blocked and water overflowed. Moreover, the food supply was probably insufficient because larvae held in that system did not grow and were not able to achieve metamorphosis: residence time of the phytoplankton was probably very low in this system. These 5 families were therefore unfortunately lost at the post-settlement stage.

Due to the inability of the first system to successfully achieve metamorphosis, a second system was set up by Dr Richard Braithwaite: an air-lift downwelling recirculation system (Utting and Spencer, 1991). Two tanks were put on top of each other, with circulation of 1 µm-filtered seawater in a closed system. The water was air-lift into the main water arrival pipe by a pump. Water level increased in the water reservoir. Then water of the reservoir overflew into each individual sieve by small pipes placed at the top of each sieve. The excess of water was finally drained off by the outer pipe. Families 10, 11, 12 and 13 were transferred into 80 µm-mesh sieves (2 sieves per family) in this system. The daily food supply consisted of a mixture of *P. lutheri*, *Chaetoceros calcitrans* (Bacillariophyceae) and *Isochrysis galbana* clone T-Iso (Prymnesiophyceae) (Figure 27b). After 20 days the larvae were transferred onto 200 µm-mesh sieves. Settled larvae were obtained from these 4 families and they were sent to Ifremer, La Tremblade (France) by DHL, aged 3 months, in plastic boxes with screw lid. There they were first held in 800 l tanks, in circular 1 mm-mesh sieves, with high input of phytoplankton to allow a good growth. They were finally transferred at age 1 year into 150 l tanks in rectangular 2 mm-mesh sieves. The height of the water in the tank was set low to avoid mixing between sieves because mussels tended to crawl to the top of the sieve and sometimes even crawled out of the sieve.

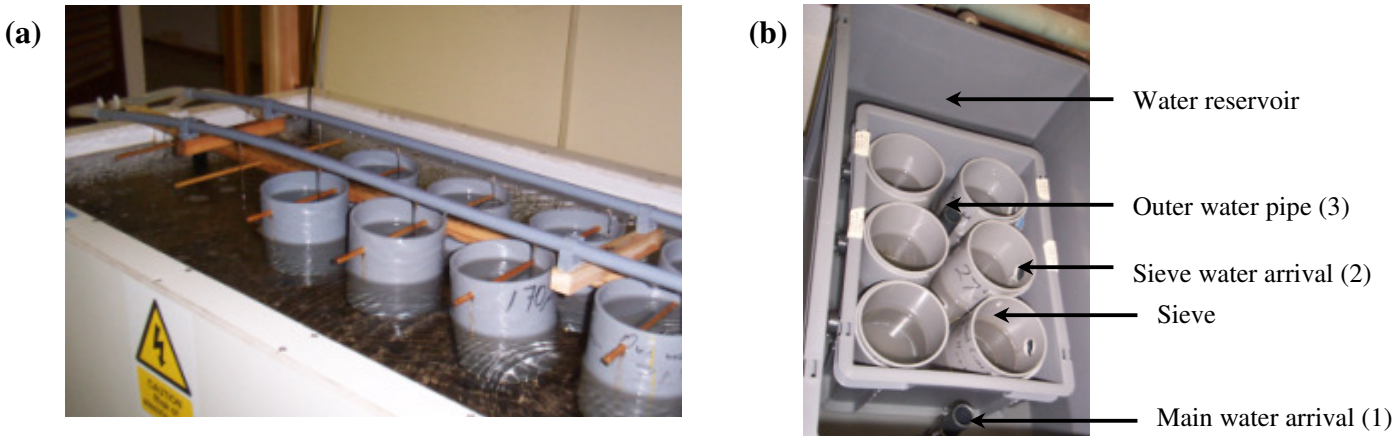


Figure 27. Post-settlement systems used. a) downwelling system, b) air-lift downwelling recirculation system. Numbers refer to the order of arrival of the water. Water is air-lift by a pump into the main water arrival.

III.2. Molecular markers: AFLPs

Thirty six AFLP primer pairs were genotyped in the mapping family: A1 to A12 (FAM fluorochrom), B1 to B12 (HEX fluorochrom) and E1 to E12 (NED fluorochrom). See Chapter 2 part III.3. for details on the AFLP methodology. The mapping family chosen, Family 10, consisted of the 2 parents (male 10 and female 10) and 92 F1 offspring. Two negative controls were included in each PCR reaction to detect any potential contamination.

III.3. Genetic mapping analysis

Construction of a genetic map requires several steps: obtaining one or several segregating families; obtaining data on linkage disequilibrium between molecular markers using parents and progeny; and linkage genetic mapping analysis itself. This analysis consists of:

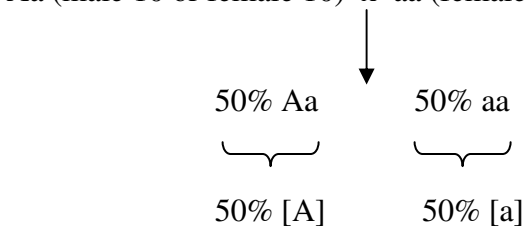
- testing segregation of each marker, i.e. testing Mendelian inheritance
- testing linkage between each pair of markers to identify linkage groups (TWO POINT analysis)
- ordering markers inside each linkage group and estimating genetic distances between them (MULTIPOINT analysis)
- testing robustness of the map by permuting neighbouring markers

Two software packages were used to establish a preliminary linkage map in *M. edulis*. MAPMAKER 3.0 (Lander et al., 1987) was used to construct sex-specific linkage maps, whereas JOINMAP 4.0 (Stam, 1993) was used to build a consensus map.

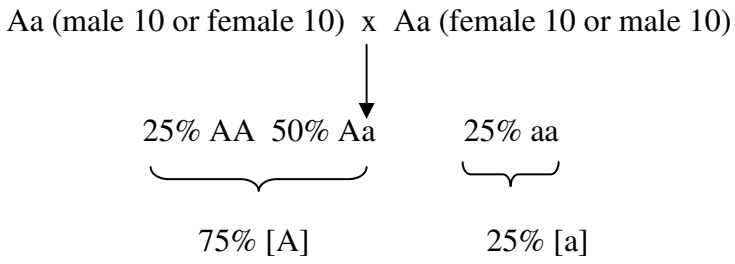
III.3.1. Segregation distortion analysis

Dataset of *M. edulis* consisted exclusively of AFLPs, which are dominant markers. Segregating markers were therefore scored for presence ([A]) or absence ([a]) of the amplified AFLP fragment. The absence of the fragment ([a]) was assumed to represent a homozygous genotype aa, whereas the presence of a band ([A]) could be one of 2 different genotypes: AA or Aa. Two kinds of segregating markers were obtained:

- type 1:1 markers corresponded to AFLPs that were present in a heterozygous state in only one of the two parents (male 10 or female 10) and were expected to segregate 1:1 in the F1 progeny, i.e. 50% of the F1 exhibiting the fragment and 50% not exhibiting the fragment



- type 3:1 corresponded to AFLPs that were present in the heterozygous state in both parents and were expected to segregate 3:1 in the progeny, i.e 75% of the F1 exhibiting the fragment and 25% not exhibiting the fragment



Type 1:1 markers were mapped in sex-specific maps, male or female because it was possible to identify which parent was heterozygous for the marker. Type 3:1 markers were mapped in both male and female maps, serving as anchor loci to build a consensus map (Figure 28).

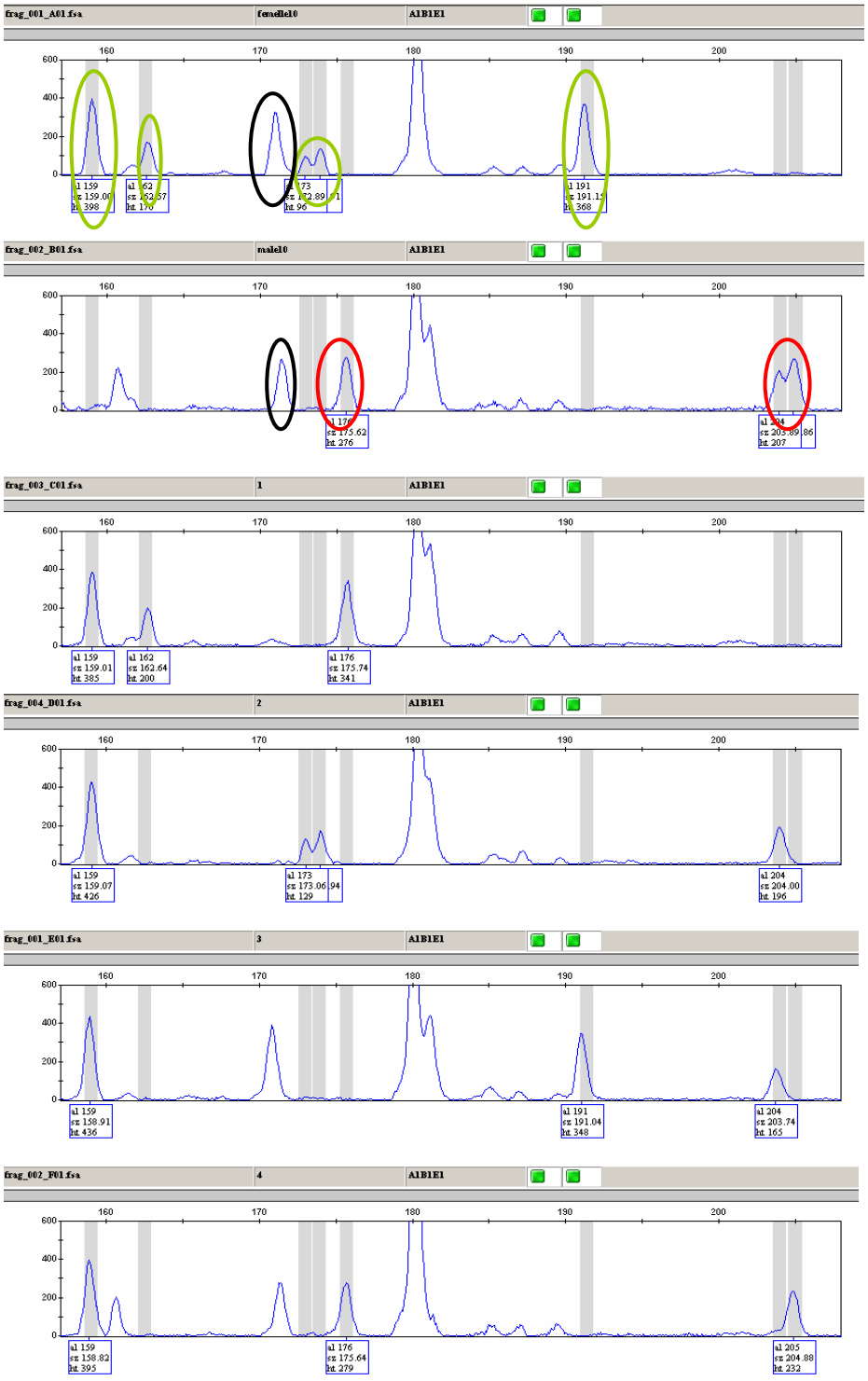


Figure 28. Electrophoregrams showing two types of segregating markers (the 1:1 type and the 3:1 type) in the *M. edulis* mapping family. Names of alleles are specified under each peak. Peak present/absent: genotype Aa/aa.

In Figure 28, 9 different markers have been highlighted, named 159, 162, 171, 173, 174, 176, 191, 204 and 205. For example, the marker 159 is present in the female (Aa), absent in the male (aa) and present in all 4 F1 progeny (Aa); this marker will be mapped in the female map only. The marker 176 is absent in the female (aa) but present in the male (Aa) and 2 progeny exhibit this fragment, F1-1 and F1-4 (Aa); this marker will be mapped in the male map only. In contrast, marker 171 (black ellipse) is present in both the female and the male (Aa) and in 2 progeny, F1-3 and F1-4 (Aa); this marker will be mapped in the male and female maps.

A χ^2 goodness of fit statistical test was applied to the 1:1 and 3:1 markers to detect segregation distortion:

$$\chi^2 = \frac{(O_{peak} - E_{peak})^2}{E_{peak}} + \frac{(O_{nopeak} - E_{nopeak})^2}{E_{nopeak}}$$

where O_{peak} and O_{nopeak} are the observed numbers of individuals with or without peak respectively; and E_{peak} and E_{nopeak} the expected number of individuals with or without peak respectively.

In the case of type 1:1 markers, $E_{peak}=E_{nopeak}=N/2$ with N the number of F1 individuals. In the case of type 3:1 markers, $E_{peak}=3N/4$ and $E_{nopeak}=N/4$. This test follows a χ^2 distribution with one degree of freedom (ν=1). The values of χ^2_{TH} are 3.841, 6.635 and 10.828 for a probability α=5%, 1% or 1‰ of making a mistake (rejecting the null hypothesis when it is true). Therefore, the following decision rule was applied:

χ^2 value	Test	Interpretation	Use for genetic mapping
$\chi^2 < 3.841$	NS	Test not significant: marker following Mendelian rules of inheritance: non distorted marker	YES
$3.841 < \chi^2 < 6.635$	p<0.05	Test significant at α=5%: distorted marker.	NO
$6.635 < \chi^2 < 10.828$	p<0.01	Test significant at α=1%: highly distorted marker.	NO
$\chi^2 > 10.828$	p<0.001	Test significant at α=1‰: very highly distorted marker.	NO

For example, for a 1:1 marker scored on 86 F1 individuals where 16 exhibited the fragment and 70 did not, the χ^2 value would be $[(16-43)^2/43]+[(70-43)^2/43]=33.9$. Therefore, this marker is very highly distorted (p<0.001) and would be discarded from further genetic linkage analysis.

III.3.2. Establishment of sex-specific linkage maps: MAPMAKER 3.0. software

MAPMAKER 3.0 software (Lander et al., 1987) was used to build sex-specific linkage maps, based on type 1:1 segregating markers. A "double pseudo-test cross" strategy was applied to our dataset: two separate data sets were obtained, one for the female parent and one for the male parent. In this pseudo-test cross configuration markers are present in one parent and absent in the other or *vice versa* (see Figure 28 for reminder) and are expected to segregate 1:1 in the F1 progeny (Grattapaglia and Sederoff, 1994). Both data sets (male and female) were analysed using the same commands (Figure 29).

Each dataset (male and female) was duplicated and recoded to allow the detection of markers linked in repulsion phase: recoded markers were adding the letter "r" at the end of their names; e.g. marker A1f123r was the recoded marker of marker A1f123. All individuals coded H for A1f123 were recoded as A for A1f123r, and vice versa. Pairs of markers for which the presence of an inverted score (absence when noninverted) was linked to presence of a noninverted score were counted as being linked in repulsion.

The initial stage of genetic mapping consists of ordering markers into linkage groups. Two markers are in linkage if the recombination fraction between them is less than 0.5. This requires a statistical test to assess if the recombination between a pair of markers is significantly less than 0.5. This hypothesis testing is often carried out by a likelihood ratio test. The LOD (Logarithm of the ODd) score is often used to assess the evidence of linkage and is defined as:

$$LOD = \log_{10} \frac{e^{L}(r = \theta)}{e^{L}(r = 0.5)}$$

where $e^{L}(r=\theta)$ is the likelihood that there is a genetic linkage between 2 markers with a recombination frequency of θ; $e^{L}(r=0.5)$ is the likelihood that the 2 markers are unlinked (r=0.5). Therefore, a LOD score of 3.0 means that the linkage hypothesis (at $r=\theta$) is 1000 ($10^{3.0}$) times more likely than the independence hypothesis. The GROUP command of MapMaker uses the two point data to determine linkage groups: pairs of markers linked with a LOD > threshold and with genetic distance < maximum distance will be placed in the same linkage group. To determine very conservative linkage groups, several criteria were defined for the GROUP command: maximum distance of 20cM, 30cM or 37.5cM; and LOD score of 2.0, 3.0, 4.0 or 10.0.

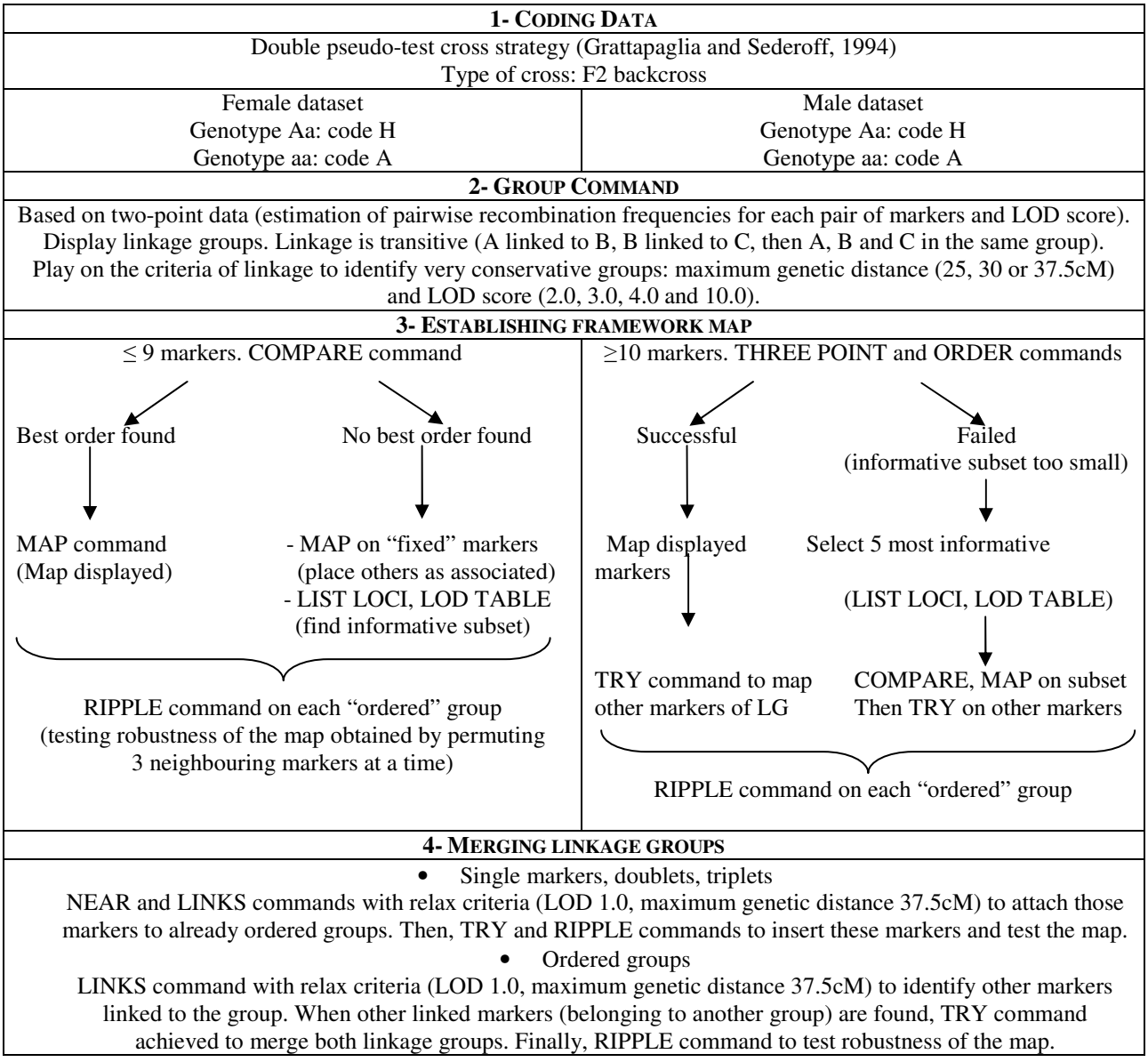


Figure 29. MAPMAKER 3.0 software: commands used for establishing sex-specific preliminary genetic maps in the blue mussel *M. edulis*, in a family issued from a single pair mating. The LOD score (Logarithm of the ODd) was used to assess evidence of linkage. A LOD score of 3.0 means that the linkage hypothesis is 1000 times more likely than the independence hypothesis. See text for further explanation.

The second stage of genetic mapping consists of ordering markers within linkage groups and estimating genetic distances between them. For this purpose, several commands can be used, depending mostly on the number of markers within a linkage group. For a limited number of markers (n≤9), the COMPARE command calculates likelihood associated with each possible

map, and displays the 20 most likely orders: this is an exhaustive search. For a higher number of markers ($n \geq 10$), this exhaustive search is too computer-time consuming. Therefore, the THREE POINT command pre-computes all 3-point data which allows the exclusion of certain orders of a set of 3 markers. The ORDER command uses the three point data directly to find the most likely map order. Because the THREE POINT command allows the exclusion of some 3 marker orders (e.g. order abc excluded based on 3 point data, order bca most likely), computation time is reduced. However, other commands of MAPMAKER can be used in the case where no most likely order is found. The LIST LOCI and LOD TABLE commands display the number of individuals genotyped for each marker and the two point data (genetic distance and associated LOD score) for each pair of markers of a given sequence of markers. These 2 functions allow the finding of an informative subset of markers, defined by markers with almost no missing data and separated by 10-20cM. This informative subset is used as a starting point for the COMPARE command to display the best order, and then other markers of the group are added one by one with the TRY command. The TRY command tests each position of a given marker in a sequence of ordered markers, and displays the LOD score associated with each position: the position with a LOD score of 0 is the most likely position of a marker and will be accepted if the second best position is associated by a LOD score of at least -2.0 (meaning that this map position is 100 times less likely). This command also displays the LOD score of the independence position (the marker is not linked to the linkage group).

The RIPPLE command allows testing of the robustness of the map obtained, by permuting neighbouring markers and displaying the LOD score of each "permuted" map. The number of markers permuted is set by the user and for this study 3 was chosen. The software displays all "permuted" maps with a LOD score < threshold (in our case 2.0) that could represent potential alternative orders. Markers that represent a conflict in map position (several map positions possible, with a small difference of LOD score) are placed as associated markers.

After ordering markers in each conservative linkage group, this is considered as a framework map and the final stage of genetic mapping consists of adding unlinked single, doublets or triplets markers, and of merging linkage groups. Clearly the ideal situation is to obtain a number of linkage groups that equals the haploid number of chromosomes (14 in the

blue mussel). Two commands were used for this purpose: NEAR and LINKS, with the relaxed criteria of a LOD of 1.0 and a maximal genetic distance of 37.5cM. For each given marker, those commands display all markers linked to them with the specified criteria. When some linkage was found, markers were added into a linkage group or 2 linkage groups were merged with the TRY command, which allows the confirmation of the linkage hypothesis versus the independence one. Finally, the new order was tested with RIPPLE command.

The ERROR DETECTION command (Lincoln and Lander, 1992) was on during all the analyses described above, to detect eventual genotyping errors. Map distances in centiMorgans were calculated using Kosambi's mapping function (Kosambi, 1944) and linkage groups were drawn with Mapchart software (Voorrips, 2002).

III.3.3. Genome length and map coverage

Average marker spacing of the framework map was calculated by dividing the total length of the map by the number of intervals (the number of framework markers minus the number of linkage groups). In the same way, the average marker spacing for each linkage group was calculated by dividing the length of each linkage group by the number of intervals on that linkage group (i.e. the number of markers minus 1).

The expected length of the genome was estimated using method 4 of Chakravarti et al. (1991): the length of each linkage group was multiplied by the factor $(m+1)/(m-1)$; m being the number of markers (framework, or framework and associated) on each linkage group. This expected genome length was first computed by taken into consideration only the framework markers but it was then re-estimated following the addition of associated markers. Observed genome coverage estimates were determined by dividing the observed genome length by the expected length of the genome. Two observed genome coverage estimates were computed, one with, and one without the associated markers taken into account.

III.3.4. JOINMAP software

Because MapMaker software cannot handle mixed segregation data (such as 1:1 and 3:1 dominant markers), the establishment of a consensus map was achieved with Joinmap 4.0

software (Van Ooijen, 2006). Male and female datasets were treated independently, as a population type CP (composite). Type 1:1 markers were coded as "lm x ll" segregation type, with presence of the fragment in the progeny coded as "lm" and absence of the fragment as "ll". Type 3:1 markers were coded as "hk x hk" segregation type, with presence of the fragment coded as "k-" and absence of the fragment as "hh". Missing values were coded as "-".

First, the two parental maps based on 1:1 and 3:1 markers were built. The classical steps of genetic mapping were applied: first grouping of markers into linkage groups based on the two-point data, and then ordering of markers inside each linkage group. For the ordering of markers, a regression mapping algorithm (Stam, 1993) was used: the software starts with the most informative pair of markers, and then adds each marker one by one. For each added locus the best position is searched by comparing the goodness-of-fit of the calculated map for each tested position. The procedure is similar to the TRY command of MapMaker. When at the best position the goodness-of-fit decreases too sharply, or when the locus gives rise to negative genetic distance estimates, this locus is removed. After adding a marker to an ordered sequence of markers, a RIPPLE command is automatically performed by the software. When at the end of the first round, all markers are not placed on the map, a second round is performed to try to add the remaining markers. Up to three rounds can be performed to order all markers of each linkage group. A map is displayed after each round. The map of the first round contains markers that have been easily and confidently ordered, whereas the third round map contains markers with ambiguous or conflicting positions. At this stage the framework markers were checked against their order in the framework map established using MapMaker.

Homologous pairs of linkage groups were identified by multiple and parallel linkages of markers, i.e several identical markers in the same order in both parental maps. Only 3:1 markers that agreed with the mapping order of framework markers (as established with MapMaker) were retained for the establishment of a consensus map.

Recombination rates were converted into genetic distances (in cM) using Kosambi's mapping function (Kosambi, 1944) and linkage groups were drawn with Mapchart software (Voorrips, 2002).

IV- RESULTS

IV.1. Segregation distortion

Among the 92 F1 progeny, 6 were discarded from further analysis because they exhibited an unexpected AFLP pattern compared with their parents. They were assumed to be contaminants from other segregating families produced at the same time in the experimental hatchery.

Thirty-six selective primer pairs were screened for 86 F1 progeny and their 2 parents. The majority of the peaks (bands) were between 60 and 400 bp. The total number of peaks, the number of segregating peaks (both types 1:1 and 3:1) and the number of distorted markers produced by each of the 36 primer pairs are shown in Table 15. All primer pairs produced identifiable peaks, ranging from 40 (pair A1) to 96 (pair A7) peaks per pair. All 36 primer pairs generated a total of 2354 peaks, averaging 65 peaks per pair. The average number of segregating markers (among the two parents, including both types of markers) was 791, or 22 per primer pair, corresponding to 33.6% of polymorphic peaks.

The total number of peaks generated or the number of segregating peaks did not vary significantly among the 3 EcoRI primers (ANOVA-one way, p=0.186 or p=0.169 respectively). However, the total number of peaks and the number of segregating markers produced varied significantly among the 12 MseI primers (ANOVA-one way, p=0.012 or p=0.003 respectively). All pairwise comparisons achieved after the ANOVA, according to Tukey's method, revealed a significant difference in the total number of peaks produced between primers Mse1 (Mse+CGA) and Mse11 (Mse+CAG). Moreover, the significant difference in the number of segregating peaks among the MseI primers was due to primer Mse2 (Mse+CAA), that generated significantly fewer segregating peaks than 6 other Mse primers (Mse1, Mse4, Mse7, Mse10, Mse11 and Mse12).

Among the 791 polymorphic markers in the mapping family (86 F1 progeny), 341 were segregating through the female parent, 296 through the male parent and 154 through both parents.

Table 15. Total number of peaks, number of segregating peaks, number of markers of type 1:1 and 3:1, and percentage of distorted markers produced by 36 AFLP primer pairs in a mapping family of *M. edulis*.
Numbers in brackets correspond to the number of distorted markers.

Primer pair	Total number of peaks	Total number of segregating peaks	Number of 1:1 markers in the female (distorted)	Number of 1:1 markers in the male (distorted)	Number of 3:1 markers (distorted)	% of distorted markers
A1	40	29	15 (2)	9 (0)	5 (2)	13.8%
B1	50	23	15 (2)	5 (4)	3 (0)	26.1%
E1	53	33	12 (3)	15 (8)	6 (2)	39.4%
A2	44	5	0 (0)	0 (0)	5 (3)	60.0%
B2	42	4	0 (0)	0 (0)	4 (2)	50.0%
E2	63	2	0 (0)	0 (0)	2 (0)	0.0%
A3	64	28	12 (4)	9 (0)	7 (2)	21.4%
B3	60	17	10 (3)	6 (1)	1 (0)	23.5%
E3	70	24	13 (2)	6 (2)	5 (2)	25.0%
A4	75	34	12 (6)	17 (5)	5 (3)	41.2%
B4	66	26	11 (3)	9 (1)	6 (2)	23.1%
E4	83	26	14 (6)	10 (4)	2 (0)	38.5%
A5	66	15	3 (1)	5 (0)	7 (3)	26.7%
B5	54	14	7 (0)	4 (4)	3 (1)	35.7%
E5	56	7	4 (1)	2 (1)	1 (0)	28.6%
A6	71	26	5 (2)	14 (5)	7 (2)	34.6%
B6	55	15	6 (2)	8 (2)	1 (0)	26.7%
E6	58	7	2 (0)	2 (0)	3 (0)	0%
A7	96	54	24 (8)	22 (7)	8 (2)	31.5%
B7	60	24	7 (5)	8 (1)	9 (1)	29.2%
E7	68	20	8 (2)	9 (2)	3 (3)	35.0%
A8	67	25	10 (2)	8 (3)	7 (0)	20.0%
B8	70	13	7 (2)	4 (0)	2 (0)	15.4%
E8	57	17	9 (0)	5 (2)	3 (1)	17.6%
A9	78	15	9 (2)	6 (2)	0 (0)	26.7%
B9	58	22	8 (1)	7 (3)	7 (2)	27.3%
E9	73	20	9 (4)	10 (3)	1 (0)	35.0%
A10	81	34	14 (4)	13 (3)	7 (4)	32.3%
B10	60	21	4 (1)	11 (4)	6 (1)	28.6%
E10	76	33	19 (7)	8 (1)	6 (0)	24.2%
A11	71	26	14 (4)	11 (5)	1 (0)	34.6%
B11	76	25	11 (2)	10 (1)	4 (1)	16.0%
E11	82	30	14 (6)	11 (3)	5 (1)	33.3%
A12	76	26	7 (3)	12 (5)	7 (1)	34.6%
B12	72	22	11 (3)	8 (1)	3 (0)	18.2%
E12	63	29	15 (5)	12 (5)	2 (1)	37.9%
Total	2354 (33.6 % polymorphic peaks)	791	341 (98)	296 (86)	154 (42)	28.6%
Total number of markers kept			243	210	112	

The percentage of distorted markers varied considerably among primer pair, ranging from 0% (e.g. E6) to 60% (A2) (Table 15). Chi-square analysis indicated that 243 (71.3%) and 210 (70.9%) markers segregated according to the expected 1:1 Mendelian ratio in the female and male respectively; and that 112 (72.7%) markers segregating according to the expected 3:1 Mendelian ratio (Figure 30). Distorted markers were discarded from further linkage analysis.

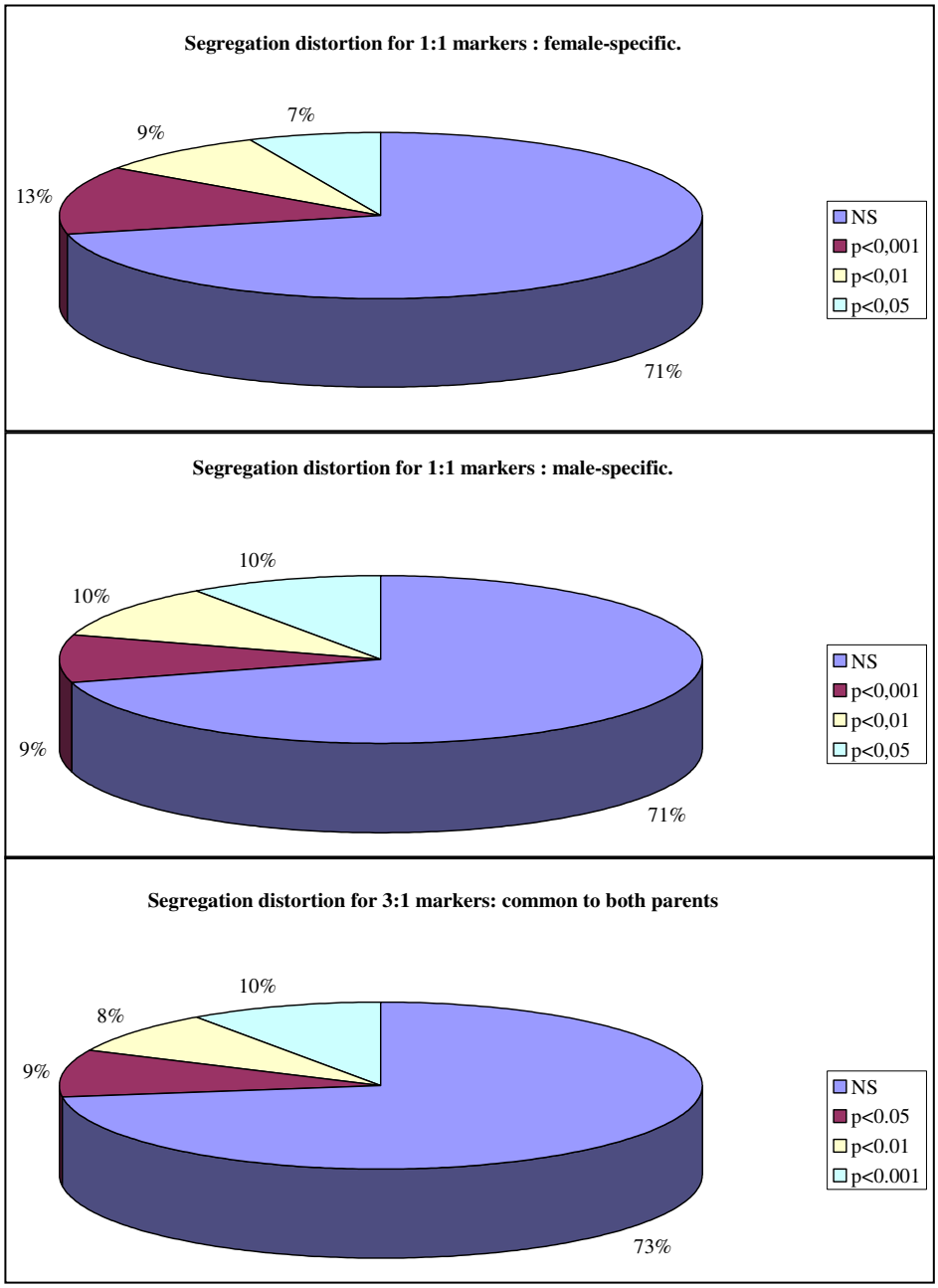


Figure 30. Segregation distortion analysis for the 1:1 and 3:1 markers of the mapping family 10 of *M. edulis*. NS: percentage of Mendelian markers; p<0.05, p<0.01, p<0.001: percentage of distorted markers at α=5%, 1% or 1‰.

IV.2. Sex-specific linkage maps

IV.2.1. Female genetic linkage map

The female framework map established with Mapmaker 3.0, based on the 243 AFLP markers segregating through the female parent only, consisted of 121 markers (49.8%). Seven markers were not linked to the framework map (2.9%). Additionally, 115 markers were linked to the framework map with a LOD score 4.0 but not placed accurately; therefore they were considered as "associated markers". Fourteen linkage groups were set up for the female map covering 862.8 cM. The sizes of the linkage groups ranged from 9.5 cM to 101.5 cM. The number of framework markers per linkage group varied from 2 to 16, and the number of associated markers from 0 to 23. The average distance between 2 framework loci ranged from 4.09 cM (G10F) to 19.5 cM (G12F), with an average spacing of 8.06 cM. The maximum interval of the female map was 32.9 cM (G2F) (Table 16).

Table 16. Length, number of markers (framework and associated), average spacing, largest interval of linkage groups of the female map.

Linkage group	Length (cM)	No. of framework markers	No. of associated markers	Average spacing (cM)	Largest interval (cM)
G1F	101.5	15	23	7.25	27.8
G2F	101.3	11	12	10.13	32.9
G3F	83.3	10	13	9.25	21.9
G4F	76.0	10	5	8.40	28.3
G5F	75.2	10	8	8.35	31.0
G6F	69.9	7	3	11.65	26.7
G7F	66.7	9	5	8.34	28.9
G8F	65.8	8	4	9.40	22.3
G9F	61.6	9	11	7.7	22.5
G10F	61.3	16	11	4.09	14.3
G11F	60.2	10	11	6.69	25.4
G12F	19.5	2	0	19.5	19.5
G13F	11.0	2	5	11.0	11.0
G14F	9.5	2	4	9.5	9.5
Total	862.8	121	115	8.06	32.9

The female framework map is displayed Figure 31. Associated markers were located beside their closest framework marker. Some clusters of AFLPs could be observed, containing from 2 to 7 markers. A single linkage group could contain up to 4 clusters (e.g. G1F).

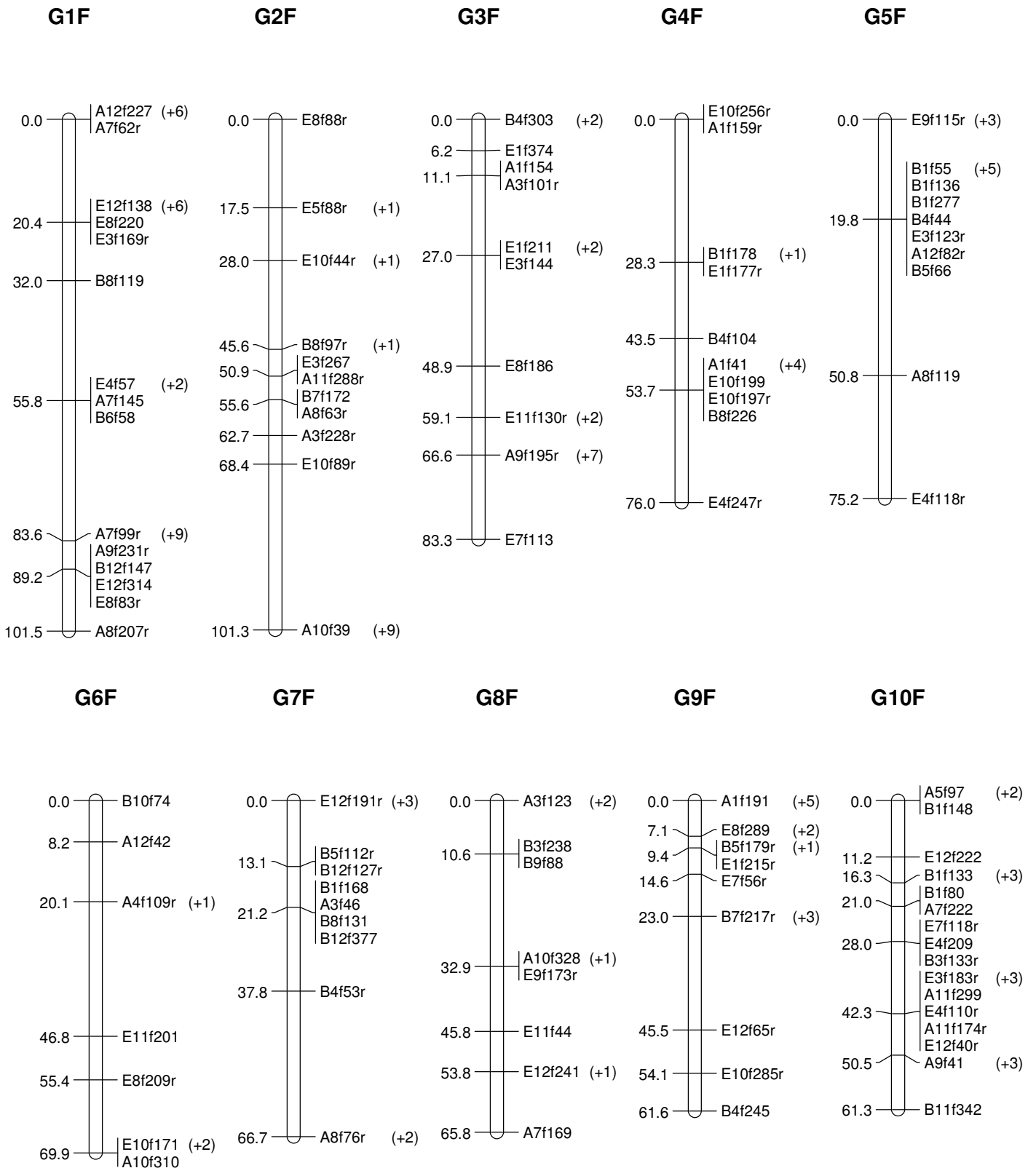


Figure 31-continued

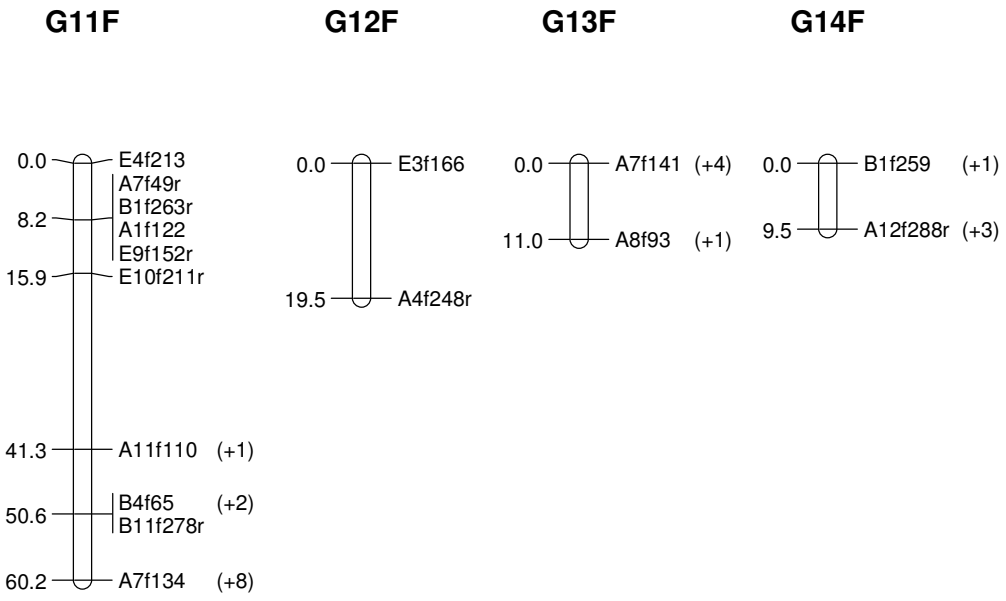


Figure 31. AFLP linkage map of the blue mussel *M. edulis*: Female map obtained with Mapmaker 3.0., 121 framework markers, 863 cM. AFLP markers are labeled with the primer pair name followed by the letter "f" (for fragment) and a 3-digit fragment size in base pairs. The letter "r" at the end of some AFLPs refers to recoded markers. Markers are indicated on the right; and absolute positions on the left (in Kosambi cM). Numbers in brackets on the right of locus name correspond to number of associated markers (linked but unplaced).

IV.2.2 Male genetic linkage map

The male framework map established with Mapmaker 3.0 was based on the 210 AFLP markers segregating through the male parent only. The resulting map consisted of 116 framework markers (55.2%). Six markers were not linked to the framework map (2.9%). Additionally, 88 markers were placed as associated markers, as they were linked to the framework map with a LOD score of 4.0 but their position could not be accurately determined with the current mapping family. Fourteen linkage groups were set up for the male map covering 825.2 cM. The sizes of the linkage groups ranged from 20.3 cM to 86.5 cM. The number of framework markers per linkage group varied from 3 to 20, and the number of associated markers from 1 to 14. The average distance between 2 framework loci ranged from 3.08 cM (G13M) to 15.4 cM (G6M), with an average spacing of 8.09 cM. The largest interval varied from 8.2 cM (G13M) to 37.6 cM (G1M) (Table 17).

Table 17. Length, number of markers (framework and associated), average spacing, largest interval of linkage groups of the male map.

Linkage group	Length (cM)	No. of framework markers	No. of associated markers	Average spacing (cM)	Largest interval (cM)
G1M	86.5	13	9	7.21	37.6
G2M	84.9	10	10	9.43	22.1
G3M	79.5	7	14	13.25	30.8
G4M	77.8	8	3	11.11	23.7
G5M	74.9	10	5	8.32	35.2
G6M	61.6	5	10	15.4	23.9
G7M	61.2	6	5	12.24	24.1
G8M	60.2	20	7	3.17	16.9
G9M	57.6	7	5	9.6	12.9
G10M	57.0	8	5	8.14	25.2
G11M	53.0	5	5	13.25	20.4
G12M	26.4	5	1	6.6	16.0
G13M	24.3	9	7	3.08	8.2
G14M	20.3	3	2	10.15	11.5
Total	825.2	116	88	8.09	37.6

The male framework map is displayed Figure 32. Associated markers were located beside their closest framework marker. As for the female map, some clusters of AFLPs could be observed, containing from 2 to 7 markers. A single linkage group could contain up to 6 clusters (G8M). All framework markers were mapped with a LOD score of 4.0 except marker A6f121r of G5M that was linked to this group with a LOD score of 2.79.

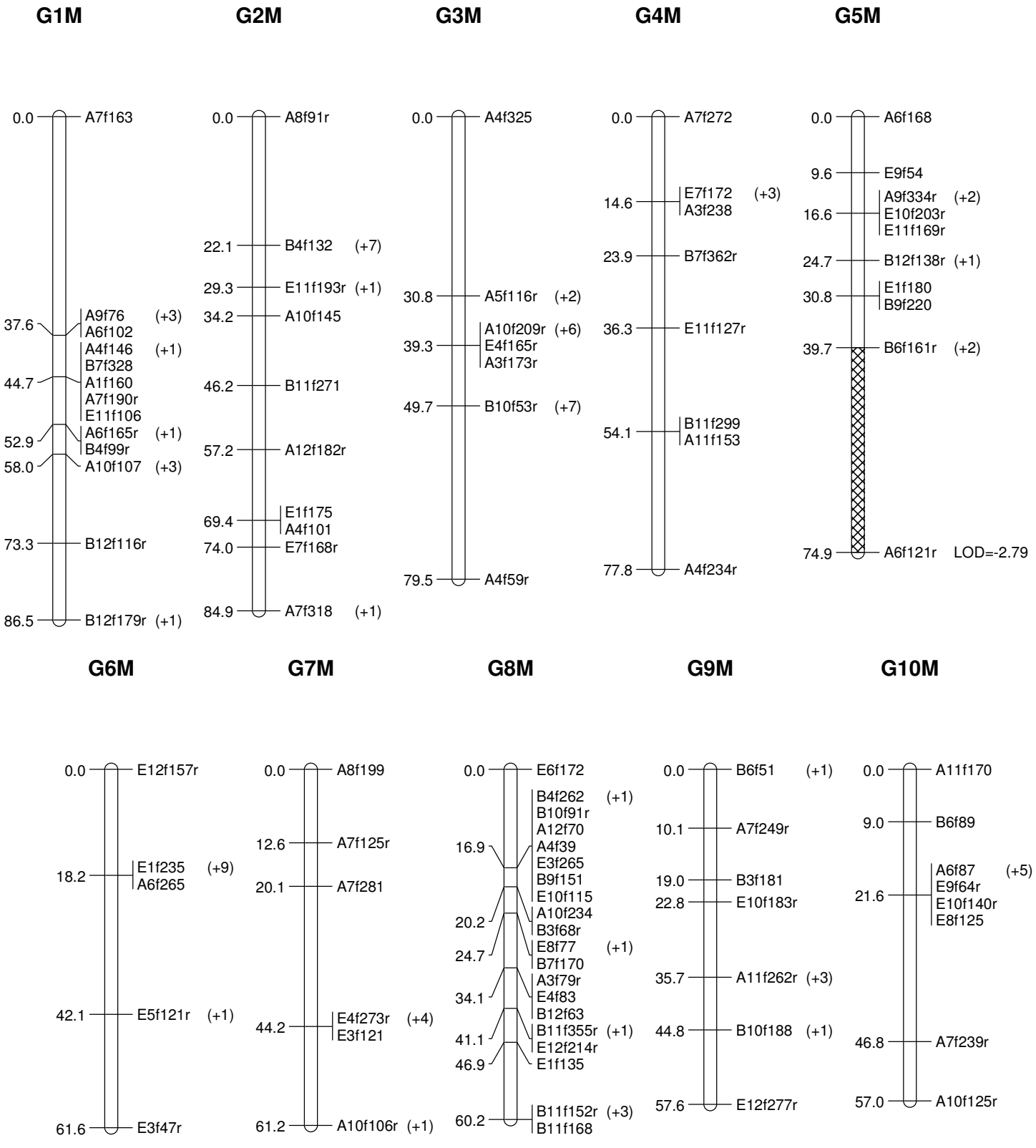


Figure 32-continued

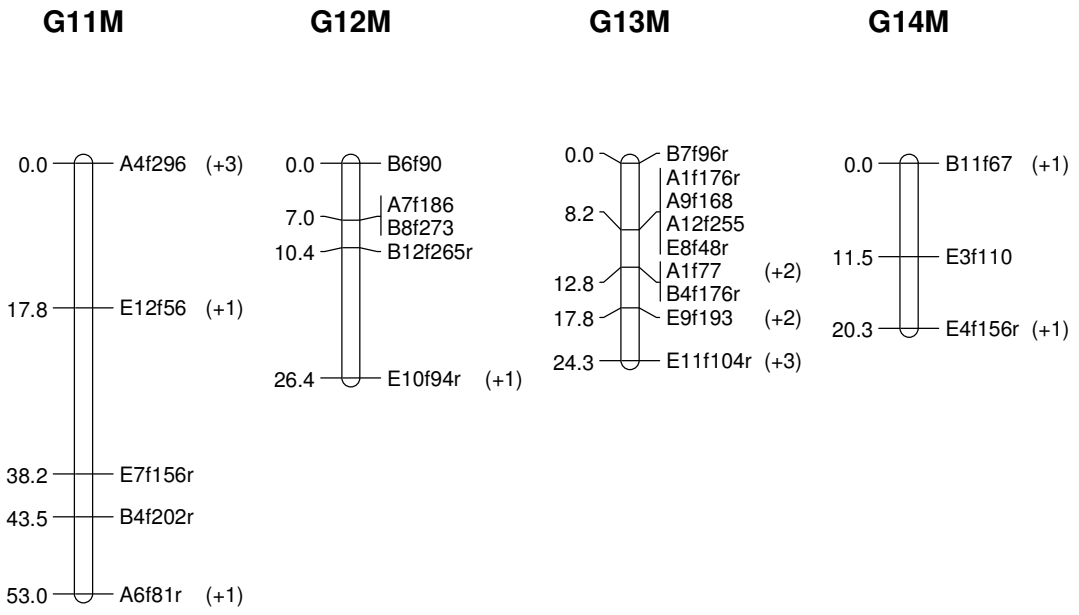


Figure 32. AFLP linkage map of the blue mussel *M. edulis*: Male map obtained with Mapmaker 3.0., 116 framework markers, 825 cM. AFLP markers are labeled with the primer pair name followed by the letter "f" (for fragment) and a 3-digit fragment size in base pairs. The letter "r" at the end of some AFLPs refers to recoded markers. Markers are indicated on the right; and absolute positions on the left (in Kosambi cM). Numbers in brackets on the right of locus name correspond to number of associated markers (linked but unplaced). Hatching on G5M refers to the placement of marker A6f121r with a LOD score<3.0.

IV.3. Marker distribution

Distributions of loci from three EcoRI primers (Eco+CAG, Eco+ACG and Eco+ACT), over the 14 linkage groups, are shown in Figure 33 for both male and female maps. In the female map, primer Eco+CAG provided 84 loci, mapping to 14 groups, ranging from 1 to 14 markers per group; primer Eco+ACG provided 72 loci, mapping to 13 groups, ranging from 0 to 13 markers per group; and primer Eco+ACT provided 80 loci, mapping to 14 groups, ranging from 1 to 11 markers per group. In the same way, in the male map, primers Eco+CAG, Eco+ACG and Eco+ACT provided respectively 88, 58 and 59 loci, mapping to 14 groups, ranging from 1 to 13 markers per group. There was a random distribution among linkage groups of markers generated by the three different EcoRI primers because a contingency chi-square test for the 14-groups x 3-EcoRI primers table was not significant for either female or male (χ^2 = 12.720 or 24.604, 22 d.f., P = 0.316 or 0.941 respectively, after pooling the last three classes). Both framework and associated markers were taken into account in those analyses.

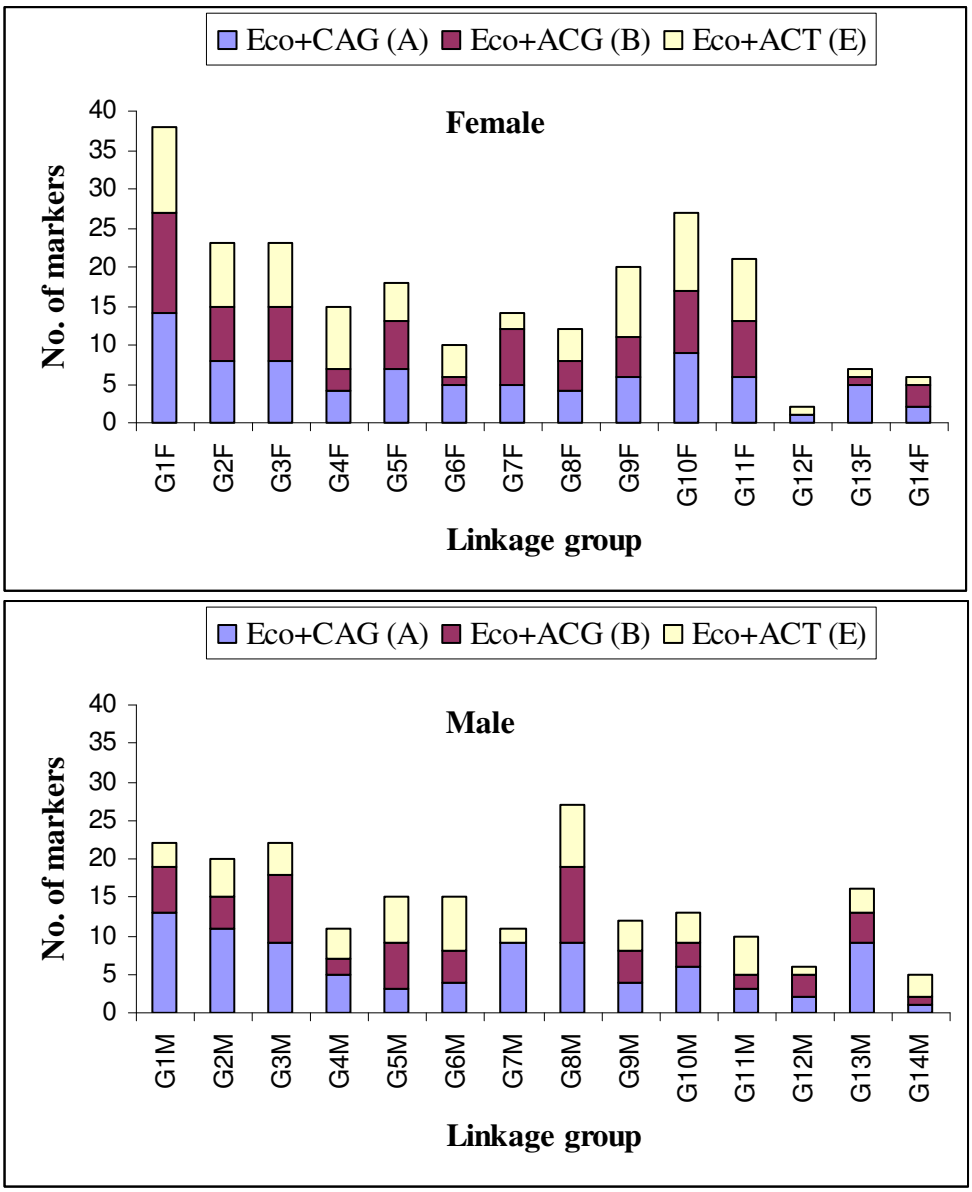


Figure 33. Histograms depicting the number of loci detected with 3 different labeled EcoRI primers, for each linkage group of the female and male genetic linkage maps of *M. edulis*. Both framework and associated markers were taken into account for establishing those distributions of markers across linkage groups.

Distribution of interval sizes between adjacent markers on the female and map framework maps revealed a high number of clusters, 43 for the female and 36 for the male. On the other hand, gaps remain to be filled as there are 16 and 12 intervals spanning more than 18 cM in the female and male maps respectively. After the cluster class, the most abundant interval length class for both maps was the one of 8 cM, corresponding to the average marker spacing (Figure 34).

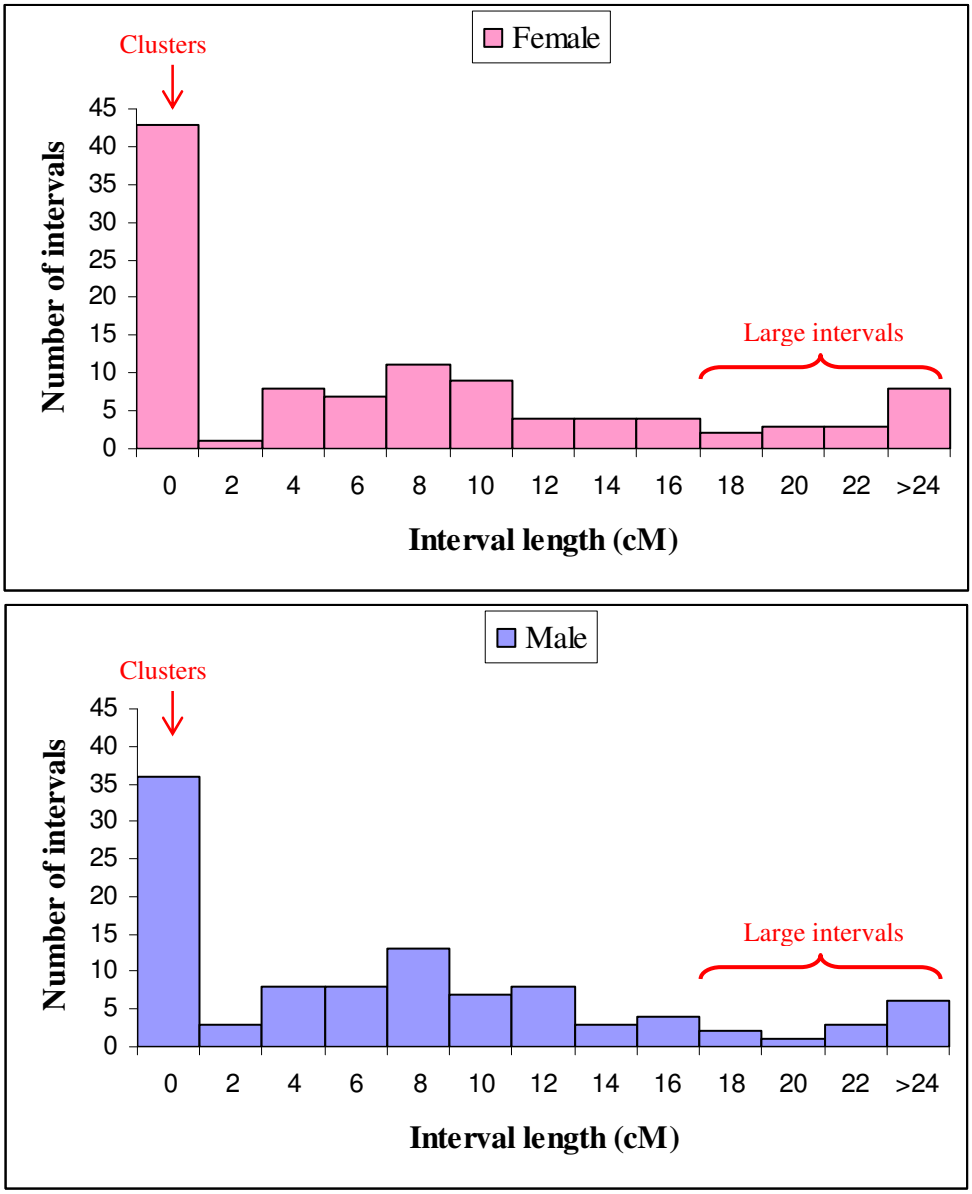


Figure 34. Distribution of interval sizes (in cM) between adjacent markers on the female and male framework maps of *M. edulis*.

The assumption of a random distribution of AFLP markers in the genome was tested in two ways: Spearman correlation coefficients and chi-square test for departure from a Poisson distribution. Spearman correlation coefficients (r_s) between genetic length and number of markers per group were 0.481 for the male ($p>0.05$) and 0.635 for the female ($p<0.05$). However, when an outlier (G8M containing 20 framework markers) was removed from the male datset, r_s increased to 0.582 and the correlation became significant ($p<0.05$). Therefore, AFLP markers tended to be

randomly distributed in the linkage maps, as indicated by the significant ($p<0.05$) correlation between the number of markers in the linkage groups and the length of the linkage groups. However, clusters were noticeable for both maps and clustering was more significant in the male map (assessed by the absence of correlation when all linkage groups are taken into account) (Figure 35).

Observed and expected distributions of AFLPs were compared for 20 cM intervals in both female and male framework maps. A chi-square test for departure from a Poisson distribution was computed. The mean of the Poisson distribution was set up to the mean number of markers per 20 cM interval length, respectively 2.61 and 2.8 for the female and male maps. No significant departure from the Poisson distribution was observed for the female (χ^2 = 13.22, 7 d.f., P = 0.067). However, this goodness-of-fit test was highly significant for the male (χ^2 = 28.63, 7 d.f., P = 0.000), mostly due to 3 intervals of 20 cM containing 8 markers. This confirms that clustering of AFLPs was more important in the male framework map, with the presence of a few high clustering of markers (containing up to 7 markers).

III.4. Genome length and genome coverage

The observed framework map length was 862.8 cM for the female map. The estimated genome length was 1125.3 cM according to method 4 of Chakravarti et al. (1991). The observed coverage was therefore 76.7% for the female framework map. When associated markers were considered, the estimated genome length was 1006.0 cM and genome coverage became 85.8% for the female map.

The observed framework map length was 825.2 cM for the male map. In the same way as for the female map, the estimated genome length was 1087.1 cM. The observed coverage was therefore 75.9% for the male framework map. When associated markers were considered, the estimated genome length was reduced to 957.6 cM and genome coverage became 86.2% for the male map.

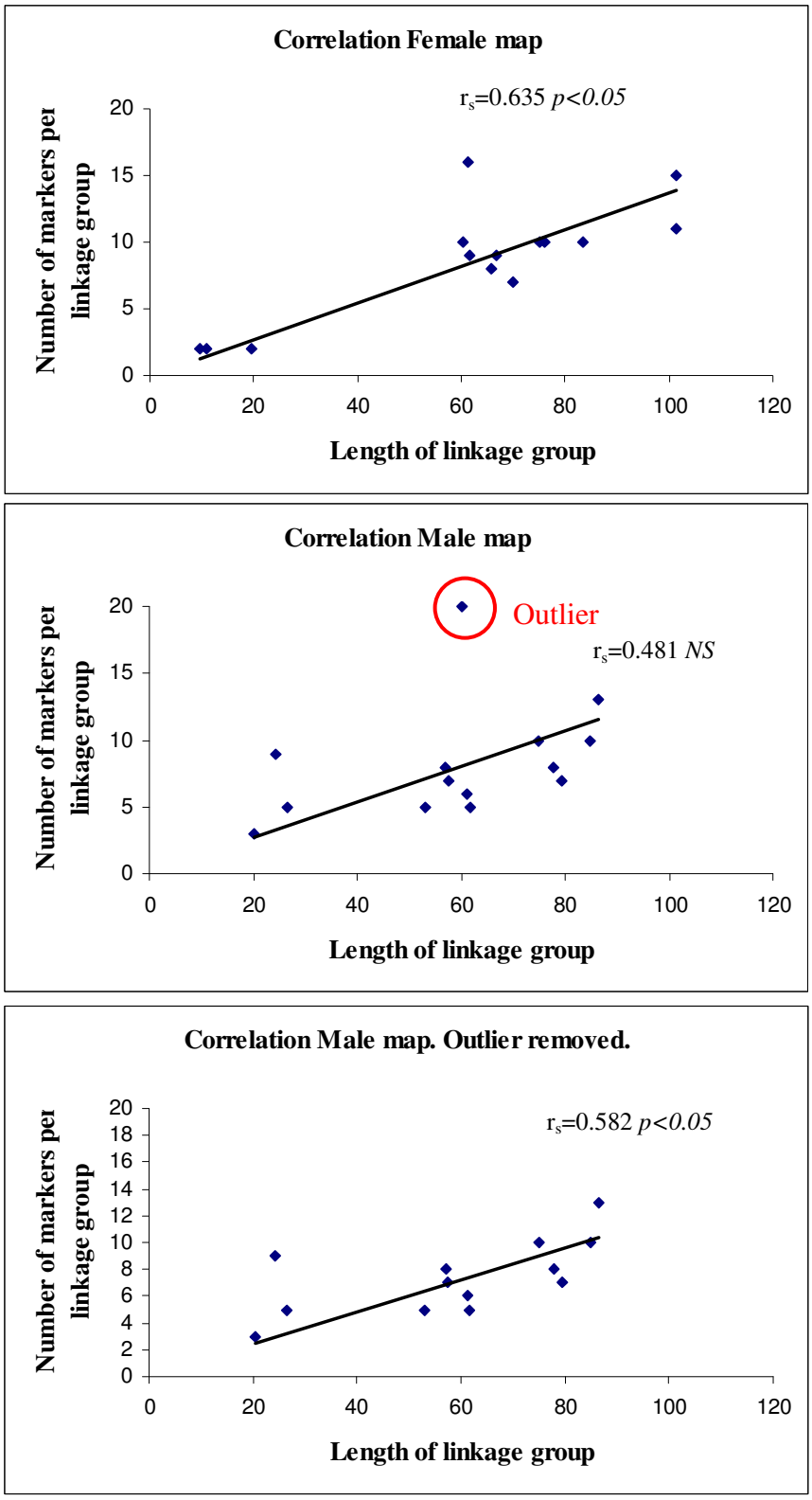


Figure 35. Correlation between genetic length of linkage group (in cM) and the number of framework markers per linkage group. Spearman correlation coefficients were assessed for the female and male maps, and for the male map after removing an outlier (G8M). *p<0.05* corresponds to a significant correlation, implying that AFLP markers were randomly distributed in the genome of *M. edulis*. *NS* corresponds to an absence of correlation.

IV.5. Towards a consensus map

The establishment of a consensus map was based on the segregation of markers of type 1:1 and 3:1. Firstly, two parental maps were constructed using the same linkage groups names as were obtained with Mapmaker. Secondly, a subset of 3:1 markers was retained that showed multiple and parallel linkages between the two parental maps without disturbing the order of framework markers established with Mapmaker.

Twelve probable homologous linkage groups were identified. For 9 of them, a consensus map was established based on at least 3 markers of type 3:1 exhibiting multiple and parallel linkages. Up to four 3:1 markers were used to build a consensus group. These consensus groups were named according to the names of the groups they derived from, e.g. consensus group G1F_G5M_comb issued from the joining of groups G1F of the female map and G5M of the male map. When homologous female and male groups were aligned, female groups seemed generally longer than their male counterparts. For example, female group G3F spanned 79 cM whereas male group G13M spanned only 32 cM. Female groups were longer than their male homologous ones in 6 out of 12 cases, of similar size in 4 out of 12 cases. In one case, the female group was far smaller than the male group: G14F spanned 21 cM, and G1M 76 cM (Figure 36). The number of pairwise comparisons between recombination frequencies of homologous pairs of markers was too low for assessing accurately any potential recombination frequency differences between sexes. More 3:1 markers per linkage groups would be needed and ideally more codominant markers to accurately compare recombination frequencies between sexes. For three of the twelve probable homologous groups, no consensus map could be built. Indeed, the homology of groups G7F-A and G14M was based on the sole marker B1f123*, which was insufficient to join both groups as at least 3 markers would be needed to orientate both groups. Moreover, the homology of groups G11F and G12M was based on the parallel and multiple linkage of 3 markers, B4f222*, E10f66* and E10f78*. However, the alignment of these 2 groups according to these 3 markers revealed that G12M could be homologous to only the terminal part of G11F, making the establishment of a consensus map difficult and not reliable. Finally, the male group G6M seemed to have two homologous in the female map: G13F and G7F-B, assessed by 3 and 2 markers of type 3:1 respectively (Figure 37).

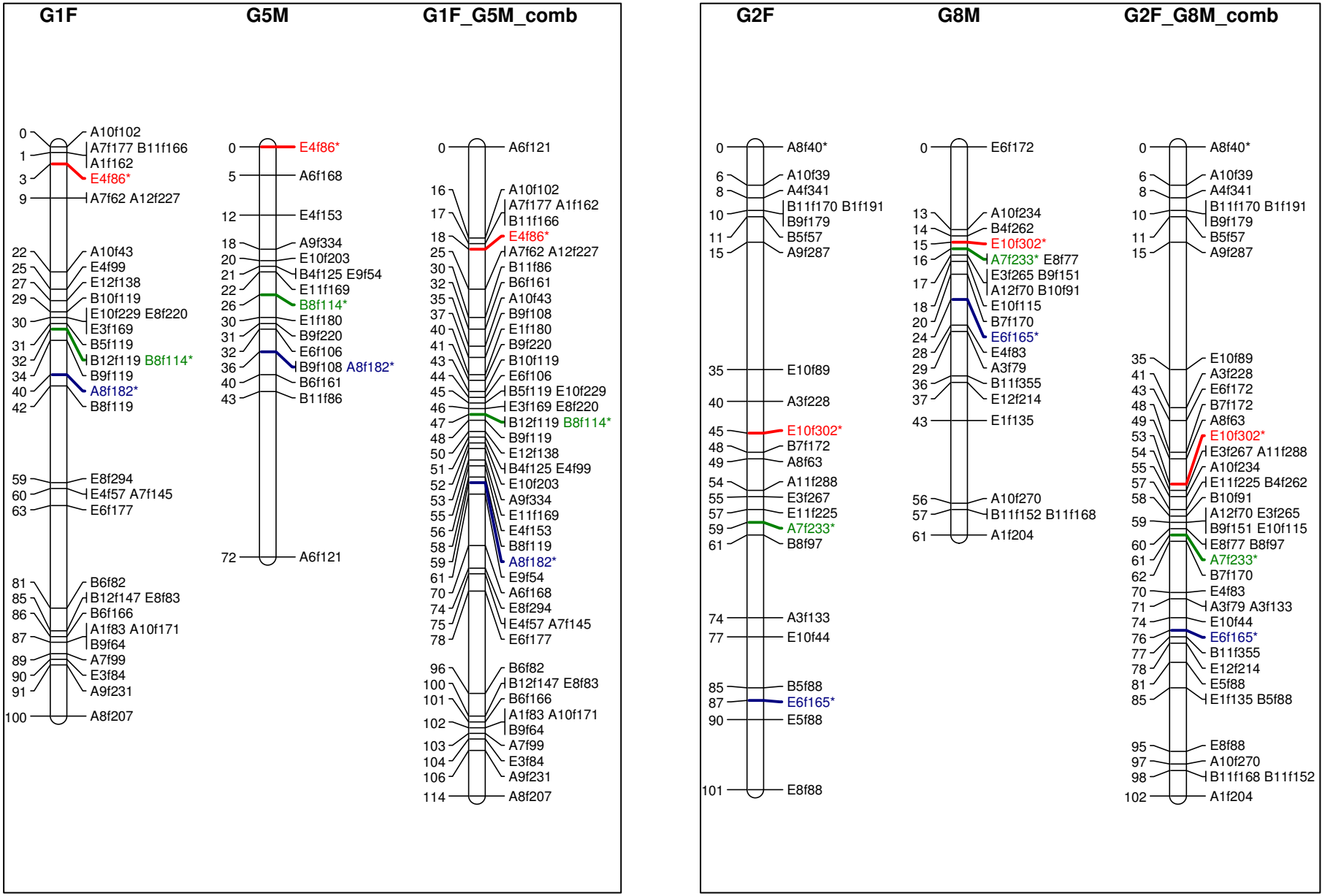
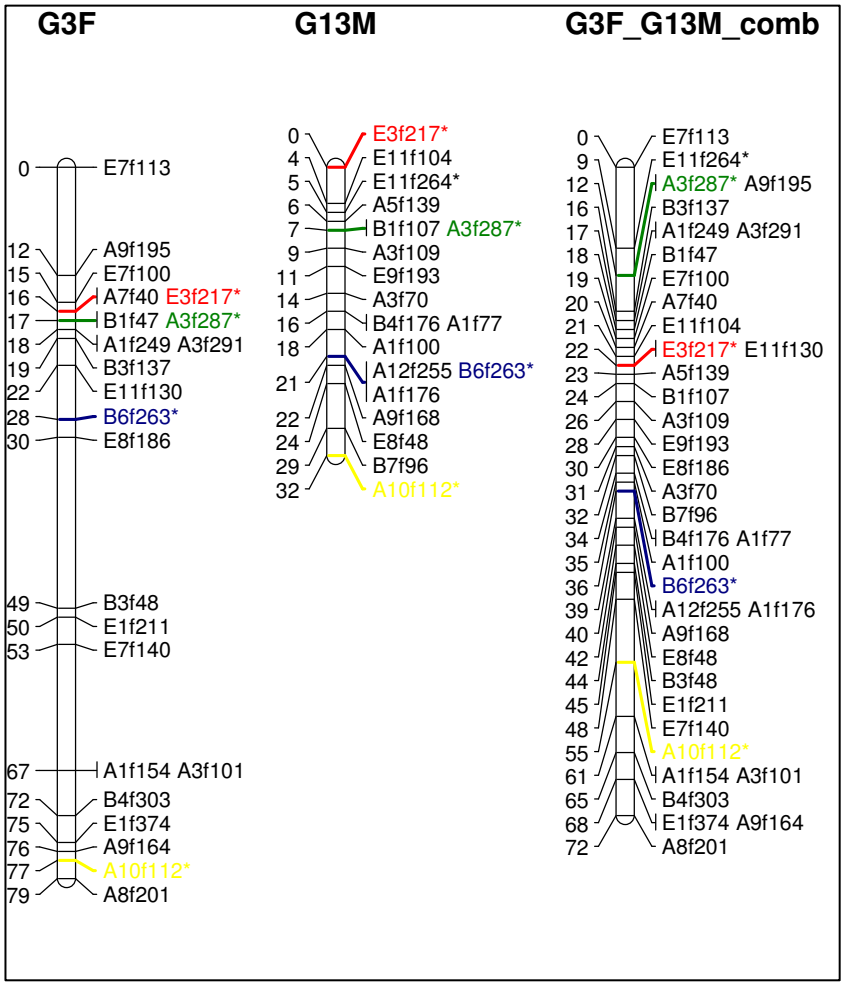


Figure 36-continued



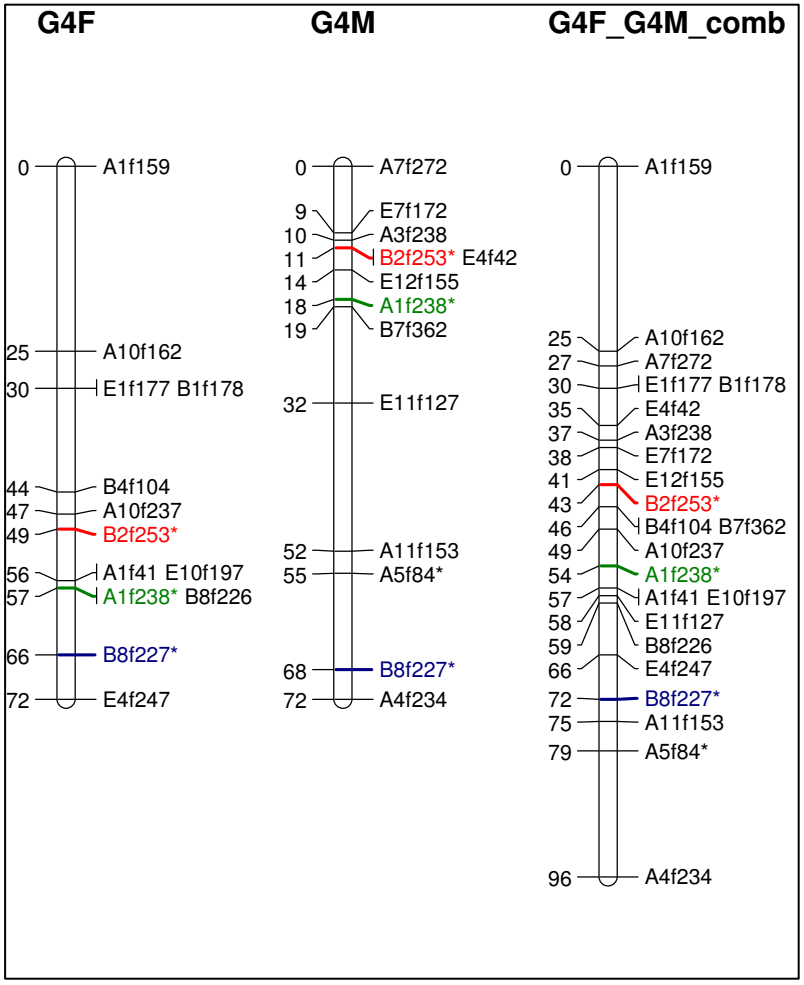
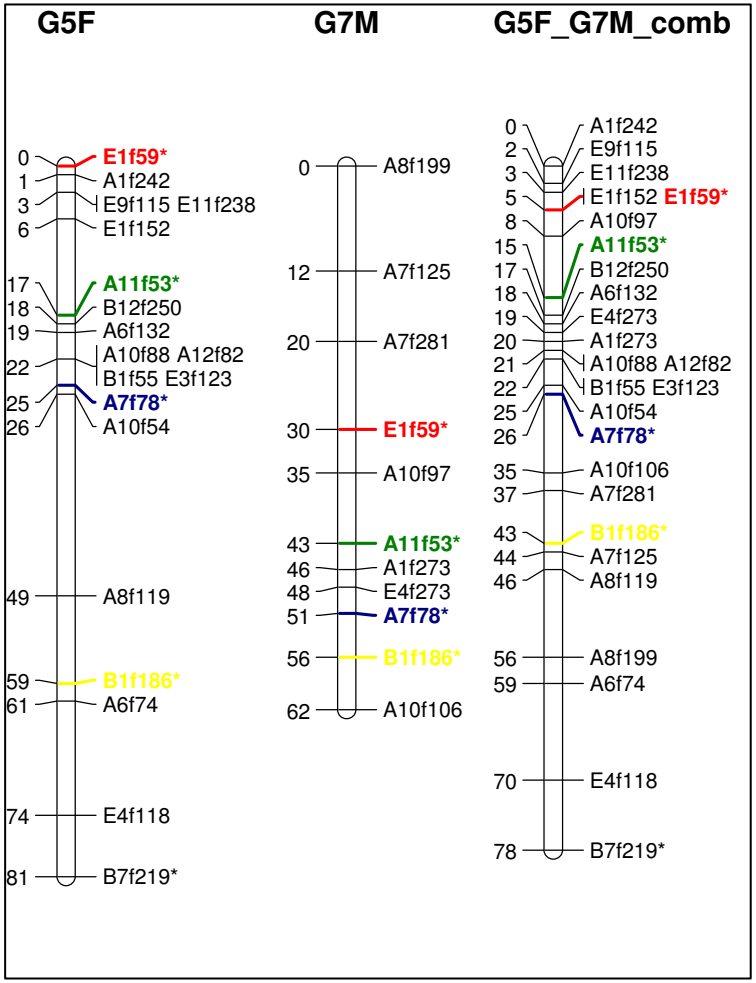


Figure 36-continued



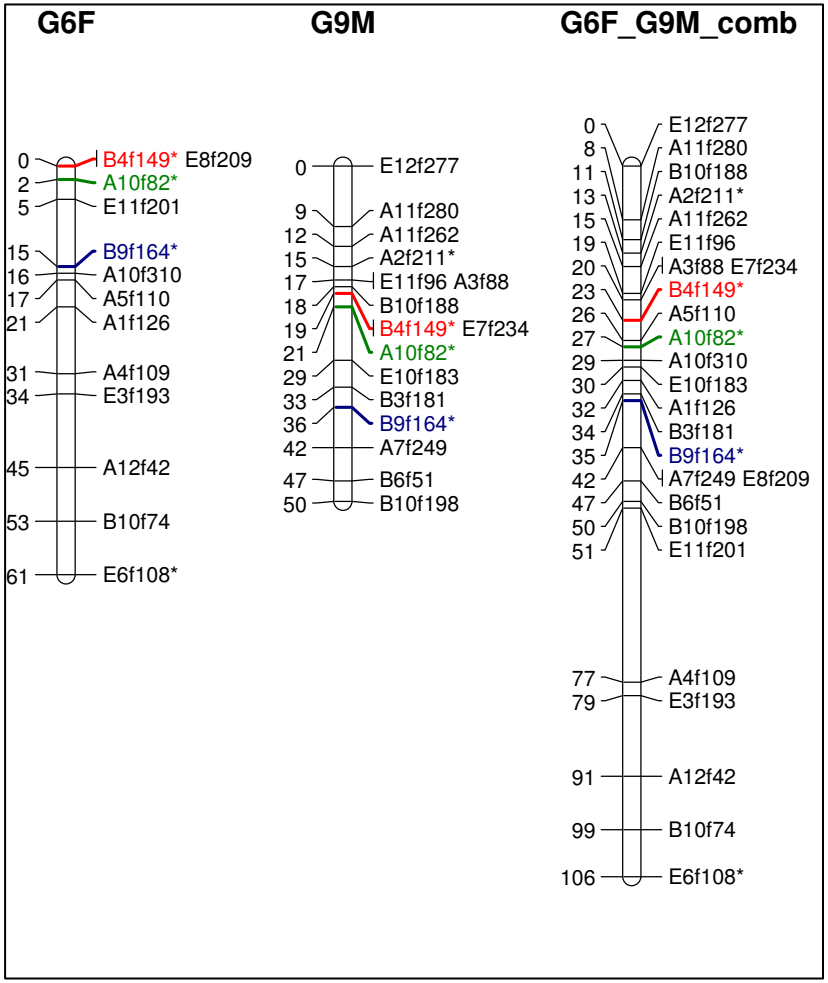


Figure 36-continued

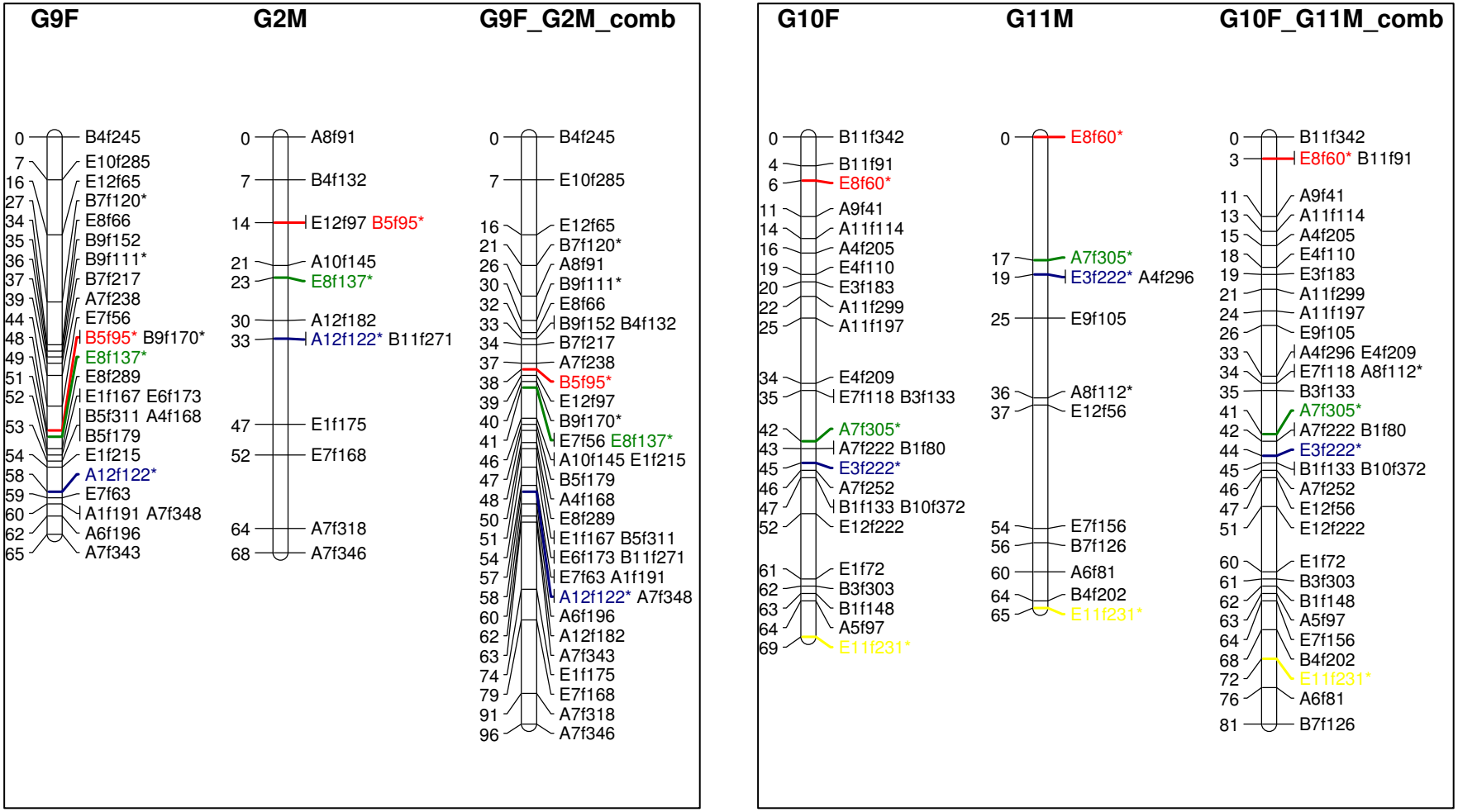


Figure 36-continued

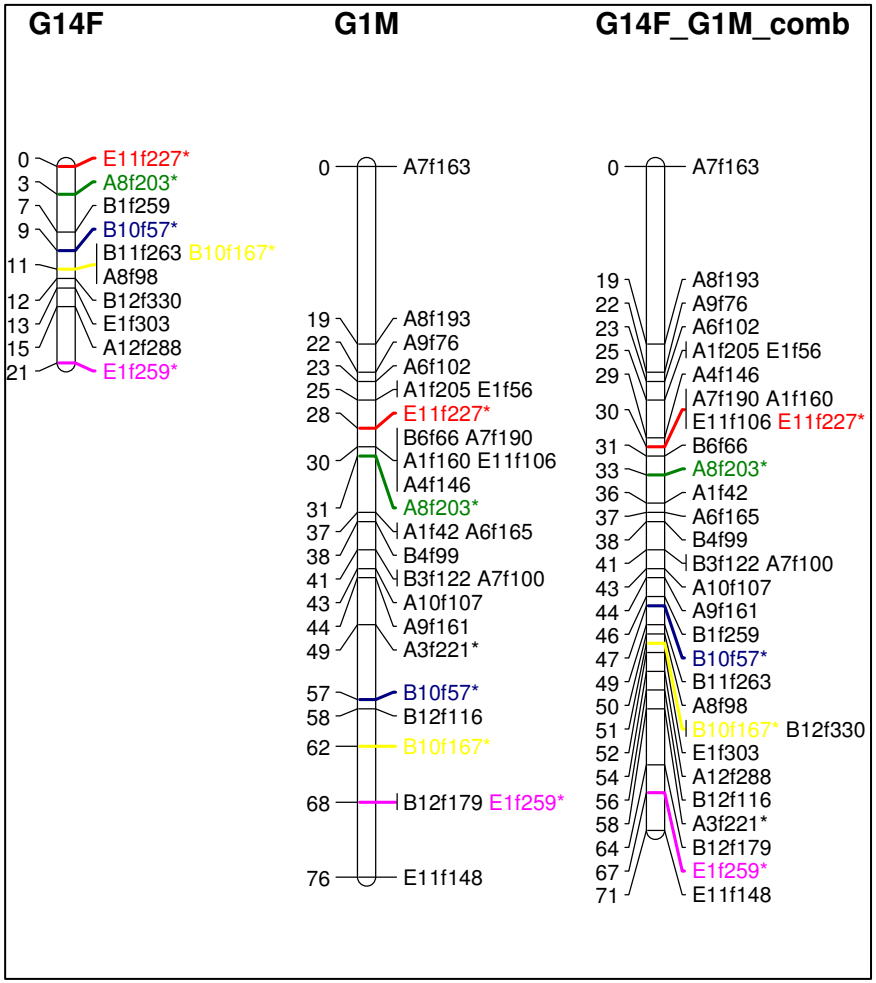


Figure 36. Consensus map established in the blue mussel *M. edulis*, in a mapping family including 2 parents and 86 F1 progeny, with Joimap 4.0 software. The consensus map presented is based on the finding of 9 homologous pairs of linkage groups. Each box refers to a single group, showing on the left the female group, in the middle the male group and on the right the consensus group. Each pair of homologous markers is displayed with a different colour to highlight multiple and parallel linkages between the two parental maps. AFLP markers are labeled with the primer pair name followed by the letter "f" (for fragment) and a 3-digit fragment size in base pairs. 3:1 markers are followed by the symbol "*". Markers are indicated on the right; and absolute positions on the left (in Kosambi cM).

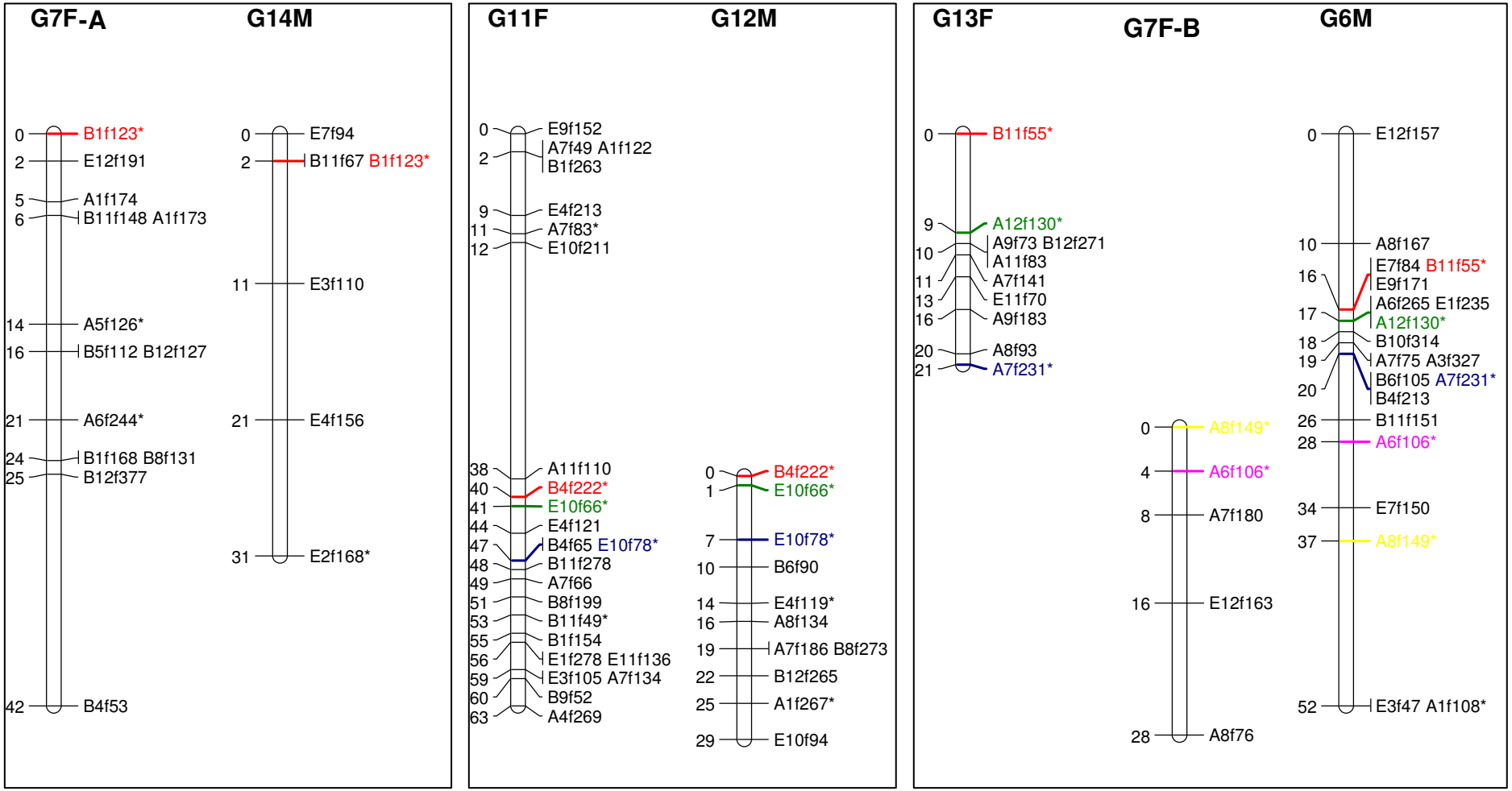


Figure 37. Probable homologies between female and male groups assessed by the mapping of 3:1 markers in *M. edulis* with Joinmap software. Three pairs of likely homologous are represented, for which no consensus map could be established. Each pair of homologous markers is displayed with a different colour. AFLP markers are labeled with the primer pair name followed by the letter "f" (for fragment) and a 3-digit fragment size in base pairs. 3:1 markers are followed by the symbol "*". Markers are indicated on the right; and absolute positions on the left (in Kosambi cM).

However, discrepancy in grouping markers into linkage groups between the two software programmes was reported for G7F. In JoinMap, this group was split into two groups, G7F-A and G7F-B according to two point data. In Mapmaker, the two terminal markers, B4f53r and A8f76r were quite further apart (around 30 cM), which could explain the splitting of G7F into two groups with JoinMap. Moreover, Mapmaker uses both maximum recombination frequency and LOD score to group markers. In contrast, JoinMap uses only one of these two criteria at once. Some discrepancy can therefore be expected. To be conservative, we chose to work on the two partial groups of G7F to find homologies.

Despite the mapping of a few 3:1 markers, no clear homologous could be identified for the male groups G3M and G10M, nor for the female groups G8F and G12F. For example, one marker of G8F was mapped in G3M and another one in G10M. The addition of new "bridge" markers is critical to resolve these ambiguities.

The total lengths obtained for the female, the male and the consensus maps established with Mapmaker and Joinmap are presented in Table 18 for comparison purposes. The consensus map built with Joinmap spanned 816 cM, distributed among 9 linkage groups. Total genetic lengths obtained with the two software were very similar.

Table 18. Total genetic lengths and number of linkage groups (in brackets) obtained for the female, male and consensus maps using two different mapping softwares, Mapmaker 3.0. and Joinmap 4.0.

	MAPMAKER	JOINMAP
Female map	863 cM (14)	871 cM (14)
Male map	825 cM (14)	799 cM (14)
Consensus map	-	816 cM (9)

V- DISCUSSION

V.1. Segregation distortion

In the mapping family of *M. edulis*, segregation distortion averaged 29% for the type 1:1 markers, and 27% for the type 3:1 markers. The observed number of distorted markers, 98 for the female 1:1 markers, 86 for the male 1:1 markers and 42 for the 3:1 markers, were higher than the expected numbers by chance only, respectively 17, 15 and 8 at α=5%. Therefore, there must be another explanation. First, segregation distortion could be caused by technical artifacts in genotyping. Size homoplasy (the fact that AFLP fragments of the same size do not belong to the same locus) could cause such distortion. Incomplete enzyme digestion and inefficient PCR could represent another technical artifact, leading to a tendency towards null homozygous AFLP genotypes (aa). In this study, around 60% of the distorted markers were deficient for homozygous null genotypes, ruling out technical artifacts as the main source of non-Mendelian segregation. Low sample size could explain a part of the segregation distortion observed but 86 F1 progeny were genotyped which is higher than some previous publications in shellfish species (e.g. 41; Moore et al., 1999) although lower than others (e.g. 102; Li et al., 2003). Finally, the relatively high proportion of distorted markers (towards a homozygous deficiency) could be explained by linkage of markers with lethal or deleterious genes in recessive state that cause genotype-dependant mortalities. This high genetic load in bivalves has been reported in several studies (Bierne et al., 1998; Launey and Hedgecock, 2001). These authors estimated the number of lethal genes as from 15 to more than 30 in *O. edulis*, and to around 12 in *Crassostrea gigas* respectively. This huge potential for genetic load in bivalve species is in accordance with the Elm-oyster model (Williams, 1975). High fecundity but with high mortality in the early stages favours individual variation in fitness and in traits correlated with fitness, leading to correlation between fitness-related traits and heterozygosity at the molecular loci (Alvarez et al., 1989; David et al., 1995; Bierne et al., 1998; Launey and Hedgecock, 2001). Evidence for this has been reported in mussels (Beaumont et al., 1995).

The range of segregation distortion reported in this study was similar to that reported for *C. gigas* in several studies: 31% (McGoldrick and Hedgecock, 1997); 20.9% (Launey and Hedgecock, 2001); 26.9% (Li and Guo, 2004). High segregation distortion was reported for Zhikong scallop *Chlamys farreri*, 37.5 % (Wang et al., 2004) and 17.8% (Li et al., 2005). However, levels of segregation distortion were lower in *Crassostrea virginica* (8.2%, Yu and Guo, 2003) and in the Pacific abalone (5.4%, Liu et al., 2006). In shrimp species, non-Mendelian segregation was far lower, zero for *Penaeus monodon* (Wilson et al., 2002), around 5% for *P. japonicus* (Li et al., 2003) or around 3% in *P. vannamei* (Perez et al., 2004).

Segregation distortion seems a common phenomenon, neither restricted to a particular type of marker nor to a particular species. This phenomenon is not restricted to bivalve species and has been reported in plants, particularly in trees such as pines, eucalypts and oaks. Barreneche et al. (1998) reported that 18% of molecular markers showed distortion in their segregation pattern and related this relatively high level of non-Mendelian segregation to the high genetic load observed in this species. Segregation distortion was also reported in several fish species, such as rainbow trout (8%, Young et al., 1998) or tilapia (13.3%, Kocher et al., 1998). Several studies based on AFLPs (Liu et al., 2003; Yu and Guo, 2003; Li and Guo, 2004), microsatellites (Barreneche et al., 1998; Young et al., 1998; Launey and Hedgecock, 2001), RAPDs (Verhaegen and Plomion, 1996), isozymes (McGoldrick and Hedgecock, 1997) reported the same phenomenon. These parallel observations support the idea of a biological origin as the main explanation for segregation distortion, i.e. high genetic load with selection against homozygous genotypes for markers linked to recessive deleterious genes.

Inclusion of distorted markers in linkage analysis may potentially distort linkage relationship by creating false linkage between markers and bias genetic distances estimations. Generally speaking, the identification of linkage assumes one basic hypothesis, Mendelian inheritance. Therefore, if this basic hypothesis does not hold, the second procedure (which tests linkage between two markers) is not applicable and the software cannot accurately estimate linkage between markers. This problem is more important with 3:1 markers. Because the ratio is already unbalanced (3/4, 1/4) for the 2 genotypes observed (A?, aa), if this ratio is distorted (towards fewer than 1/4 of individuals lacking the fragment), the software lacks the information

to accurately estimate recombination frequencies. These estimates may therefore be biased and could cause disturbance in the ordering of non-distorted markers. Therefore, it is critical to discard all distorted 3:1 markers from any linkage analysis study. Nevertheless, slightly distorted 1:1 markers ($p<0.05$) could be added to a framework map (based exlusively on the Mendelian markers) in a second stage, on the condition that they do not disturb order of existing markers.

One way to reduce or avoid segregation distortion is to score molecular markers at the larval stage, before early larval mortalities occur. This strategy was used by Hubert and Hedgecock (2004) and proved to be very efficient in reducing distortion of segregation ratios. The main inconvenience of this approach is the low DNA concentration extracted from larvae which hampers the genotyping of hundreds of markers in a single larval sample. Therefore, scoring more genetic markers (in a subset of the progeny assay) and discarding distorted markers in an adult or juvenile mapping family is still a very valuable strategy.

Finally, the mapping of distorted markers may help to understand the distribution of deleterious recessive genes in the genome. Indeed, Yu and Guo (2003) reported the clustering of 6 distorted markers in an area spanning 5 cM, which could potentially correspond to the presence of a deleterious gene nearby. In the same way, Li and Guo (2004) reported the mapping of at least 4 major deleterious recessive genes in the female map of the Pacific oyster, highlighted by the clustering of markers with segregation distortion in the same direction. In the rainbow trout map, large regions of linkage groups contained blocks of distorted markers that could be linked to sub-lethal genes (Young et al., 1998).

V.2. Efficiency of AFLP methodology for genetic mapping, coupled with pseudo-test cross strategy

This study showed that AFLP methodology is an efficient way of generating relatively quickly and at a reasonable coast a high number of molecular markers that can be used for genetic mapping purposes. Indeed, this study reported an overall rate of polymorphism of 33.6%, averaging 22 markers per primer pair. This level of polymorphism could have been increased by pre-screening the 2 parents and a subset of their progeny, to select the most polymorphic primer

combinations. However, even without pre-screening, this level of polymorphism was similar to the one reported by Yu and Guo (2003), Li and Guo (2004) in *C. gigas* or higher than the one reported in shrimp species (Wilson et al., 2002; Li et al., 2003; Perez et al., 2004). The use of AFLPs is all the more important in species for which scarce genetic resources are available because they require no preliminary knowledge of the genome and can saturate genetic maps relatively quickly. Therefore, this technology was critical to the success of establishing a preliminary genetic linkage map in the blue mussel *M. edulis* because only 7 microsatellites have been published (Presa et al., 2002) and, although many allozyme markers have been developed for the mussel, low polymorphism makes them unsuitable for extensive mapping (Beaumont, 1994a). In shellfish, AFLPs represent so far the most commonly used markers and have been used to construct preliminary maps in various shrimp species (Moore et al., 1999; Wilson et al., 2002; Li et al., 2003), in the American oyster (Yu and Guo, 2003) and in the cupped oyster (Li and Guo, 2004). In fish species, AFLP-based genetic maps were established for tilapia (Kocher et al., 1998; Agresti et al., 2000), rainbow trout (Young et al., 1998), walking catfish (Poompuang and Na-Nakorn, 2004), Japanese flounder (Coimbra et al., 2003) and medaka (Naruse et al., 2000).

However, despite their wide use in genetic mapping studies, AFLPs remain a poor alternative when other markers are available. Their main limitation is that they are dominant markers, and are therefore less informative than codominant markers. However, by adopting a pseudo-test cross strategy, AFLPs become as informative as codominant markers as there are no more genotypes missing. The pseudo-test cross strategy was developed by Grattapaglia and Sederoff (1994) in an Eucalypt species. It relies on the selection of single-dose polymorphic markers that are heterozygous in one parent, null in the other and therefore are expected to segregate 1:1 in their progeny as in a testcross. This strategy, coupled with the AFLP technology, is a very powerful tool for the establishment of genetic maps, and has been employed in a variety of species, from shellfish (Li et al., 2003; Yu and Guo, 2003; Li and Guo, 2004; Perez et al., 2004) to plant species (Verhaegen and Plomion, 1996; Barreneche et al., 1998).

V.3. Linkage map, map length and genome coverage

To our knowledge, maps presented in this study represent the first genetic linkage maps established in the blue mussel *M. edulis*. Both maps, male and female, despite their preliminary nature, appear to offer a good representation of the blue mussel genome. Firstly, both maps contain 14 linkage groups, which correspond to the haploid number of chromosome of this species (Thiriot-Quievreux, 1984a).

Secondly, total map length observed in this study is similar to the theoretical genetic length based on 1.0-1.3 crossing over per chromosome. The observed genetic length was 825.2 cM for the male map, and 862.8 cM for the female map. Work on the Pacific and Eastern oysters revealed an average number of chiasmata per chromosome of ~1.1-1.2 (Guo, X., Yang, H. and Wamg, Z., unpublished data cited in Li and Guo, 2004). Based on these data, assuming a hypothetical range of 1.0-1.3 chiasmata per chromosome for *M. edulis*, the theoretical map length should range from 700 to 910 cM (1.0 or 1.3 x 50 cM x 14 chromosomes). The observed total genetic length for both maps therefore falls into that range. Moreover, expected genome lengths estimated in this study (957-1006 cM) are 5-36% longer than the theoretical length based on cytological studies (700-910 cM). Discrepancy between both estimates (expected and theoretical genome lengths) is far smaller than the one obtained for *C. gigas*: 42-99% (Hubert and Hedgecock, 2004); 32-79% (Li and Guo, 2004).

Thirdly, the ratios of longest to shortest linkage groups are respectively 10.7:1 and 4.3:1 in the female and male maps. These ratios are greater than the cytological ratio (length of chromosome 1 to length of chromosome 14) observed in several karyological studies in *M. edulis*: 2:1 (Thiriot-Quievreux, 1984a); 1.74-1.86:1 (Insua et al., 1994). This confirms the fact that gaps remain to be filled and that more markers should be added for a better coverage of the genome. However, Hubert and Hedgecock (2004) reported slightly higher ratios, 7.6:1 in the female and 13.7:1 in the male.

Fourthly, genome coverage estimated for both maps is relatively good: 76.7% and 75.9% for framework female and male maps respectively. Genome coverage increases to around 86%

for both maps when associated markers are taken into account. These estimates are similar to the one established in *C. gigas*: 70-79% (Hubert and Hedgecock, 2004) or 81-92% (Li and Guo, 2004); and in *C. virginica*: 70-84% (Yu and Guo, 2003). Another criterion reveals that framework maps established in *M. edulis* cover a good proportion of the genome: only 3% of markers are unlinked to any other marker in both maps.

Recombination patterns may vary across the genome. Indeed, even if markers were globally randomly distributed across the genome (Figure 35), some clustering of AFLPs has been observed as well as places where long intervals occur between them (Figure 34).

Clustering of AFLPs was mostly observed in the male map where up to 7 markers could be mapped at a single position of the genome, and up to 6 of these clusters were detected on a single linkage group (G8M). Clustering of AFLPs is a common feature of AFLP-based genetic maps and has been observed in several species, e.g. in maize (Castiglioni et al., 1999), rainbow trout (Young et al., 1998), tilapia (Agresti et al., 2000) or channel catfish (Liu et al., 2003). Clustering of AFLPs could result from the non-random distribution of enzymatic restriction sites across the genome, and therefore indirectly from the choice of enzymes of restriction used. In this study, EcoRI and MseI were used for digesting the DNA. As EcoRI and MseI restriction sites are relatively AT-rich, clustering observed could reflect the variation in GC content among chromosomal regions (Yu and Guo, 2003). In maize, the use of the restriction enzymes PstI/MseI resulted in more evenly distributed AFLP markers, possibly because of the combimation of an AT-rich site (MseI) and a relatively GC-rich one (PstI) (Castiglioni et al., 1999). Similarly, the use of 2 combinations (PstI/MseI and EcoRI/MseI) in a shrimp species resulted in the absence of clustering of AFLPs (Wilson et al., 2002). Contrary to studies on oysters, clustering of AFLPs seems to occur more towards centromeric regions than telomeric ones in the blue mussel. This observation is particularly obvious in groups G1M, G3M or G10M. As chromosomes of *M. edulis* are mostly metacentric or submetacentric (Thiriot-Quievreux, 1984a; Insua et al., 1994), this clustering could therefore correspond to centromeric suppression of recombination, associated with heterochromatin (Tanksley et al., 1992).

Large intervals were observed in both maps, with a maximum interval between framework markers of 32.9 cM (female) or 37.6 cM (male). Large intervals have been observed in previous shellfish studies (e.g. Hubert and Hedgecock, 2004; Li and Guo, 2004; Wang et al., 2004), fish species (e.g. Young et al., 1998) or plant species (e.g. Plomion and O'Malley, 1996; Barreneche et al., 1998). These large intervals could be due to the medium-density of the maps obtained, and it is expected that adding markers should reduce those gaps. Alternatively, they could correspond to hot-spot regions of recombination in the genome.

Physical mapping coupled with genetic mapping has allowed the visualisation of recombination rates throughout the length of the chromosomes in several plant species (Shah and Hassan, 2005). Recombination hot spot regions have been identified in wheat (Faris et al., 2000) and maize (Brown and Sundaresan, 1991). Suppression of recombination around the centromere and/or Nucleolar organizer regions (NORs) has been demonstrated in rice (Cheng et al., 2001), 2 species of grass (King et al., 2002) and in Hessian fly (Behura et al., 2004). Therefore physical mapping has shown that clustering of markers and large intervals may not only be due to artifacts of genetic mapping but could also result from different dynamics of recombination during meiosis, due to regional structural differences across the genome (e.g. euchromatin versus heterochromatin). Unfortunately, physical mapping has not yet been achieved in mussels.

Physical versus genetic sizes are reported in Table 19 for 4 species of bivalves. Genome size (or C-value) corresponds to the haploid nuclear DNA content. For *M. edulis*, this C-value has been estimated to be 1.6 pg (Hinegardner, 1974) or 1.71 pg (Rodríguez-Juíz et al., 1996). DNA content in pg can be converted in physical genome size in base pairs, 1 pg corresponding to 978 Mbp (http://www.genomesize.com). Therefore, physical genome size of the blue mussel is estimated to approximately range between 1565 and 1672 Mbp. Thus the physical to genetic size ratio is estimated as 1.56-1.75 Mb/cM for *M. edulis*. Ratios were calculated for 3 other bivalve species (*C. gigas*, *C. virginica* and *O. edulis*) in Table 19. Physical size varies by a factor of 2 between *C. virginica* and *M. edulis*. *Ostrea edulis* shows the smallest genetic length, almost half the size of *M. edulis*. Moreover, the ratio of physical to genetic size ranges from around 0.5 Mb/cM (*C. virginica*) to 2.2 Mb/cM (*O. edulis*). The higher this ratio, the more genes a candidate area of the genome will contain. Therefore, this ratio is of fundamental importance to determine

the potential for mapping a QTL down to a candidate gene. The candidate-gene approach should be more easily implemented in a species like *C. virginica* because it appears to have a smaller physical size per cM

Table 19. Genome features (genetic and physical) of 4 species of bivalves. n: haploid chromosome number; Mb: mega base pair (10^9 base pair); cM: centiMorgan.
1: Hinegardner, 1974); 2: Rodríguez-Juíz et al., 1996); 3: Gonzalez-Tizon et al., 2000); a: this study; b: Li and Guo, 2004); c: Hubert and Hedgecock, 2004); d: Yu and Guo, 2003).

Species	Physical size (Mb)	Expected genetic length (cM)	n	Mean genetic size per chromosome (cM)	Physical/genetic size ratio (Mb/cM)
M. edulis	1565-1672 (1,2)	957-1006 (a)	14	68-72	1.56-1.75
C. gigas	890 (3)	758-1030 (b)	10	76-103	0.86-1.2
		776-1020 (c)	10	78-102	0.87-1.1
C. virginica	675 (1)	858-1296 (d)	10	86-130	0.52-0.79
O. edulis	1144 (2)	513-532 (a, chapter 4)	10	51-53	2.1-2.2

V.4. Limitations of the study and future work

The main limitation of this study is that maps established were based solely on AFLP markers, which are dominant. High repeatability of AFLPs was assessed in this study, as well as by Jones et al. (1998) across 7 laboratories. Portability of AFLPs across laboratories does not seem to be a problem, at least as long as AFLP genotypes are collected with the same scoring systems from similar automatic genetic analysers. However, transferability of AFLP markers from one family to another mapping family is not certain, and will need to be assessed. It will probably depend on the level of polymorphism of the species studied, and on the degree of relationship between the two families studied. This could therefore hamper the portability of the genetic maps established.

Another limitation of this study is that the number of F1 progeny scored are rather too few to accurately map markers that are closely linked to each other in the genome. Indeed, mapping markers differing by 1 cM needs the identification of one recombinant among 100 individuals. Our dataset, with 86 F1 progeny, is therefore unable to hit such accuracy. This explains why the framework maps contained 55.2% and 49.8% of markers in the male and female maps

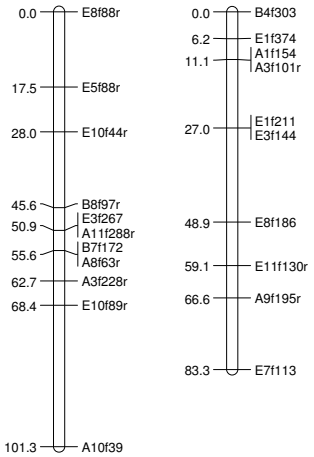
respectively, the others being identified as associated markers and placed beside their closest framework marker.

The consensus map established was based on the segregation of 3:1 AFLP markers, i.e. common in both parents and segregating in the F1 progeny assay. However, those types of markers are more difficult to map because they are less informative and exhibit missing data (the presence of a fragment is associated to 2 genotypes, AA or Aa). This means that recombination fractions between 1:1 and 3:1 markers, or between two 3:1 markers are estimated with less accuracy (Ritter et al., 1990). Therefore, to identify homologous pairs of linkage groups, multiple and parallel linkages of markers were retained for establishing the consensus map while 3:1 markers that disturbed the order of framework markers were discarded from the analysis. The same strategy has been successfully applied in the maritime pine: a two-way pseudo-test cross mapping strategy; construction of 2 parental maps with MapMaker and then with JoinMap; resulting in the construction of an AFLP-based consensus map with JoinMap (Chagné et al., 2002). To our knowledge, no AFLP consensus map has yet been published for a shellfish species. The consensus map presented in this study, even though incomplete, shows the feasibility of such a strategy. However, to increase accuracy of the consensus map, more 3:1 markers should be scored, to base the consensus map on more than 3 markers per group, and markers should be chosen to be more evenly spaced throughout each linkage group. More importantly, adding codominant markers such as microsatellites or SNPs (higher information content), serving as anchor loci between the 2 parental maps, will increase the accuracy of the consensus map built as well as its portability. Sex-averaged microsatellite genetic maps have been built in several fish species, such as channel catfish (Waldbieser et al., 2001) or tilapia (Knapik et al., 1998; Shimoda et al., 1999). Moreover, the combination of both types of markers, dominant (AFLPs or RAPDs) and codominant (microsatellites), proved to be very useful for the construction of a consensus map in rainbow trout (Nichols et al., 2003), tilapia (Kocher et al., 1998), zebrafish (Johnson et al., 1996) or common carp (Sun and Liang, 2004).

In conclusion, the framework maps presented in this study, although preliminary, achieved a cover of more than 75% of the blue mussel genome, with an average of one marker per 10 cM. According to Darvasi et al. (1993), the power of detecting a QTL is good for an average marker spacing of 10-20 cM. Therefore, the maps produced could be used for a QTL mapping approach. However, to increase the portability of any QTLs found, codominant markers such as microsatellites, SNPs or ESTs need first to be developed and added to the framework maps. The addition of such new markers should help to improve the accuracy of the consensus map and assist with analyzing sex-specific differences in recombination rates.

CHAPTER 4

Genetic mapping in *Ostrea edulis*



I- INTRODUCTION

Two main groups can be distinguished among oysters: the cupped oysters (including two genera: *Crassostrea* and *Saccostrea*) with a deep cupped shell and the flat oysters (including the genus *Ostrea*) with a flat upper shell.

The flat or "native" oyster, *Ostrea edulis*, is endemic to the Atlantic and Mediterranean coasts of Europe. Naturalised populations are also found in eastern North America from Maine to Rhode Island, following intentional introductions in the 1940's and 1950's (Jaziri, 1990), and in California. *Ostrea edulis* exhibits interesting reproductive characteristics such as sequential protandrous hermaphroditism with possibility of changing sex several times in the same reproductive season and brooding of eggs and early larvae in the mantle cavity (Yonge, 1960; Le Dantec and Marteil, 1976).

The flat oyster industry was of considerable economical importance in the 19th century in France and Britain (Neild, 1995). Massive mortalities occurred around the turn of the 20th Century from which the industry never recovered. Oyster aquaculture production fell further from 30 kt in 1970 to 6 kt today because of two successive parasitic diseases, marteiliasis (due to *Marteilia refringens*) and bonamiasis (due to *Bonamia ostreae*).

To date, genetic studies in the flat oyster have dealt mainly with spatial genetic structure (population genetics), genetic load, temporal genetic structure or cytological studies. Population genetic studies using allozymes failed to detect fine-scale genetic differentiation in European stocks of *O. edulis* (Johanesson et al., 1989; Saavedra et al, 1993; 1995). Microsatellites and mitochondrial 12S-rDNA marker have been used to investigate genetic differentiation among populations at the European geographic scale (Launey et al., 2002; Diaz-Almela et al., 2004), among North Atlantic populations (Sobolewska and Beaumont, 2005; Beaumont et al., 2006) or within North American populations (Vercaemer et al., 2006). Two regions were reported as genetically differentiated, Atlantic-western Mediterranean and eastern Mediterranean (Launey et al., 2002; Diaz-Almela et al., 2004). Moreover, a high genetic load as well as the effect of inbreeding on growth and survival in the early developmental stages was reported in *O. edulis*

(Bierne et al., 1998). Other studies focused on the temporal genetic structure with estimation of effective population size in natural and experimental populations (Diaz-Almela et al., 2004; Hedgecock et al., 2007). Finally, cytological studies revealed that the diploid chromosome number in *O. edulis* is 2n=20 (Longwell and Stiles, 1973; Thiriot-Quiévreux and Ayraud, 1982; Thiriot-Quiévreux, 1984b) and karyotype studies have shown a variable number of NORs (Nucleolar Organiser Regions) on several chromosomes (Thiriot-Quiévreux and Insua, 1992). So far, no banding technique (C-banding or G-banding) has been employed on this species for the characterisation of heterochromatin patterns but the diploid DNA content of *O. edulis* was estimated by flow cytometry to be 2.33 ± 0.02 pg (Rodríguez-Juíz et al., 1996).

Genetic linkage maps consist of ordering molecular markers across the genome based on the estimation of recombination frequencies in structured controlled populations. Linkage maps have been developed in an increasing number of aquaculturally important fish or shellfish species such as salmon (Moen et al., 2004b), tilapia (Kocher et al., 1998) and shrimp (Li et al., 2006b). Such maps represent a framework which enables the identification and localisation of QTLs (Quantitative Trait Loci) for traits of interest, such as growth or resistance to a disease, with the final aim of achieving genetic improvement through marker-assisted selection (MAS) (Liu and Cordes, 2004). Genetic linkage maps have been established in several commercially important bivalves: the Pacific oyster *Crassostrea gigas* (Hubert and Hedgecock, 2004; Li and Guo, 2004), the Eastern oyster *Crassostrea virginica* (Yu and Guo, 2003) and the Zhikong scallop *Chlamys farreri* (Wang et al., 2004; Li et al., 2005; Wang et al., 2005). QTLs have been mapped only recently in a few shellfish species, *C. virginica* (Yu and Guo, 2006) and the Kuruma prawn (Li et al., 2006a). Data on bivalves are scarce and no genetic map has yet been constructed in any flat oyster species.

Therefore, the aim of this part of the research was to:

- produce segregating families with known parental and grand-parental genotypes
- choose one mapping family based on DNA polymorphism
- establish a first genetic linkage map on that species, based on microsatellites and AFLP markers

II- BIOLOGY OF *OSTREA EDULIS*

II.1. Taxonomy

O. edulis is a bivalve mollusc that has an oval or pear-shaped shell with a rough, scaly surface, with concentric sculpture and fine radiating ridges. The two valves of the shell have different shapes: the left valve is concave and fixed to the substratum, the right being flat and sitting inside the left. The shell is off-white, yellowish or cream in colour with light brown or bluish concentric bands on the right valve. Hinge line does not exhibit teeth in the adult. The inner surfaces are pearly, white or bluish-grey, often with darker blue areas. The single adductor muscle scar is white, or slightly discoloured. This species grows up to 110 mm long, rarely larger (Howson and Picton, 1997)

Taxonomy of the European flat oyster *O. edulis* (Mollusc, Bivalve, Filibranchia) is shown in Table 20.

Table 20. Taxonomy of the European flat oyster *O. edulis*.

TAXONOMY	DESCRIPTIVE FEATURES
Phylum: Molluscs	Soft body, unsegmented, often protected by a calcareous shell, and comprising the mantle, the visceral mass, the foot and the mantle cavity.
Class: Bivalves (or Lamellibranchia)	Shell with two valves Gills arranged in lamellae No head Clear bilateral symetry
Superorder: Filibranchia	Four demibranchs (2 inner, 2 outer), each composed of descending and ascending filaments of lamellae (gill plicae)
Family: Ostreideae	Left valve of the shell deeper than the right one One adductor muscle
Genus: Ostrea	Shell of the adult irregularly oval, with valves having the same external outline, the upper valve being slightly sculptured by radiating ribs
Species: edulis	Pear-shaped shell, concentric structure and fine radiating ridges Hinge line without teeth in the adult

II.2. Geographic distribution and habitat

The genus *Ostrea* has a very wide distribution, in warm and temperate waters of all oceans. The species *O. edulis*, mostly restricted to temperate climate, therefore distinguishes itself from the other species of the *Ostrea* genus that have generally a more tropical distribution.

The European flat oyster *O. edulis* is a species endemic from European coasts and can be found along the western European coast from Norway to Morocco in the north-eastern Atlantic and in the whole Mediterranean Basin (Figure 38). Natural populations are also observed in eastern North America from Maine to Rhode Island, following intentional introductions in the 1940s and 1950s (Jaziri, 1990).

O. edulis is associated with highly productive estuarine and shallow coastal water habitats on firm bottoms of mud, rocks, muddy sand, muddy gravel with shells and hard silt.

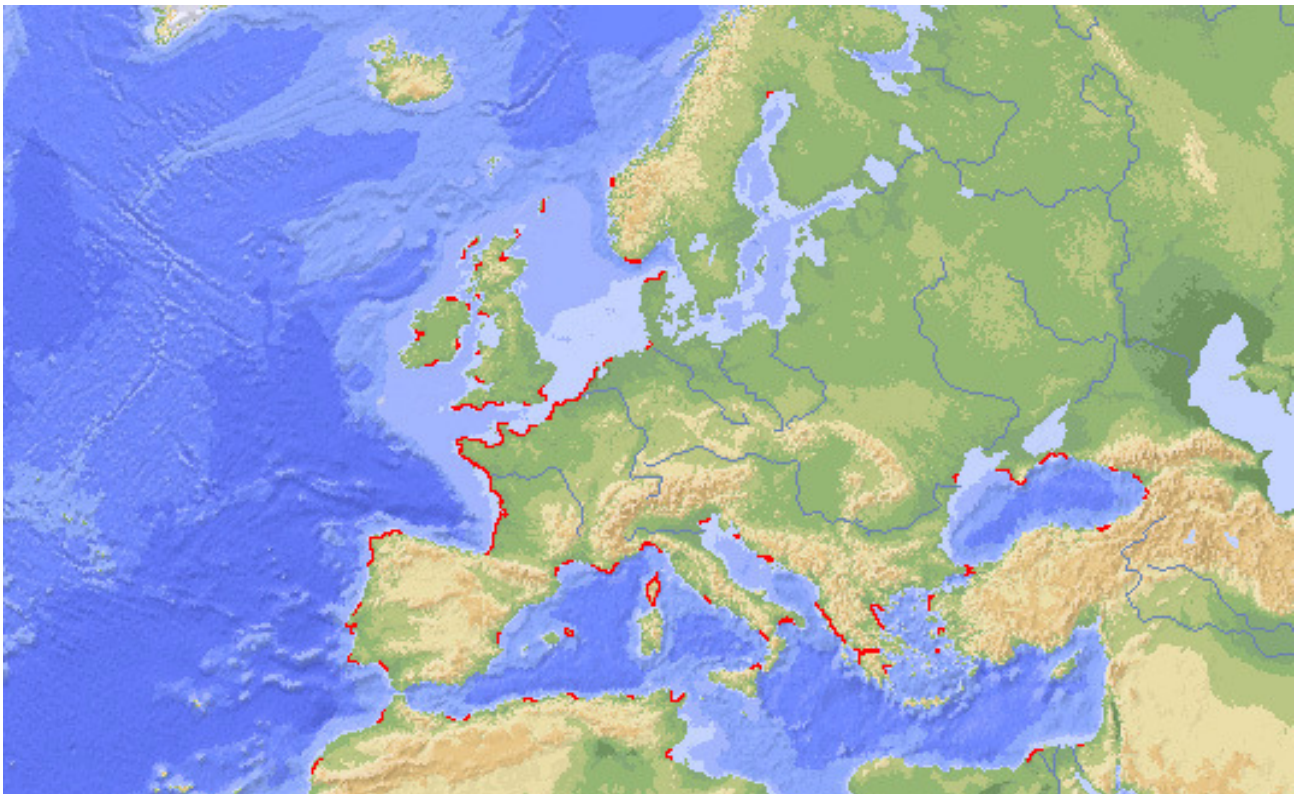


Figure 38. Geographic distribution of *O. edulis* (red edging) (from Jaziri, 1985)

O. edulis is an active suspension feeder. Feeding is carried out by pumping water through the gill chamber, where suspended organic particles are collected by gill cilia and passed to the mouth in mucus strings.

II.3. Reproductive cycle

Flat oysters are protandrous alternating hermaphrodites. This means that when first mature at 3 years, they start off as males producing sperm then switch to egg-producing females, back to males and so on. Flat oysters are able to change of sexes several times in the same reproductive season (Le Dantec and Marteil, 1976).

Gamete maturation begins in March or April and is in part temperature dependent. Gametogenesis may be continuous in warmer conditions. There may be some periodicity in spawning with peaks during full moon periods. Fecundity may be as high as 2 million eggs in large individuals (Wilson and Simons, 1985).

The main reproductive feature of the flat oyster in comparison with the cupped oyster is that this species is brooding. The eggs are around 150 µm in diameter. Eggs produced during the female stage are held in the gills and mantle cavity. The eggs are fertilised by sperm filtered with the inhalant water flow used for feeding and respiration. The fertilised eggs are retained for 7-10 days in the mantle cavity of the female during the veliger stage. Then larvae are released into the water column, around 1 million per spawning, at a size of 160 µm, and spend 5 to 14 days as a pelagic stage before settlement, depending on the water temperature (Yonge, 1960). Larval growth and survival rates are obtained in salinities of 20‰, although they can survive at salinities as low as 15‰. Growth is quite rapid for the first year and a half and then slows down after five years (Le Dantec and Marteil, 1976) (Figure 39). Market size is from about 70-100 grammes.

A life span of 5-10 years is probably typical (Quero and Vayne, 1998).

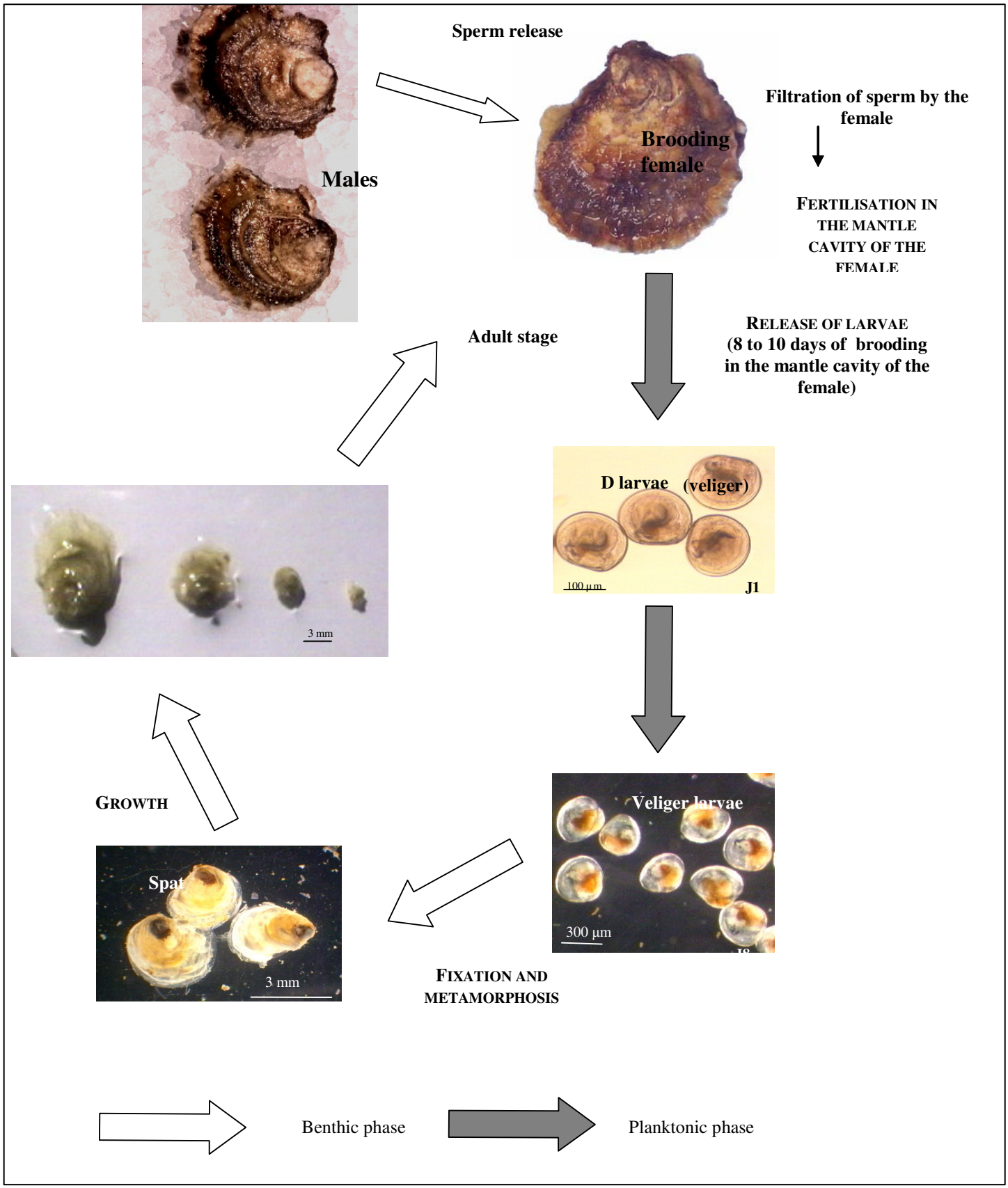


Figure 39. Reproduction cycle of the European flat oyster *O. edulis* (from Launey, 1998)

II.4. Cytological features

The first cytological studies in *O. edulis* consisted of the establishment of a karyotype by GIEMSA-staining. It allowed the characterisation of gross morphology of chromosomes, i.e. estimation of chromosome number and classification of chromosomes into classes (metacentric, submetacentric, subtelocentric, telocentric) according to the position of their centromere. All studies agreed for a diploid chromosome number of 2n=20 (Longwell and Stiles, 1973; Thiriot-Quiévreux and Ayraud, 1982; Thiriot-Quiévreux, 1984b). Both studies of the 1980's reported 5 pairs of metacentric (pairs 1 to 5), 2 pairs of submetacentric (pairs 6 and 7), 2 pairs of submetacentric or subtelocentric (pairs 8 and 9) and one pair of small submetacentric (pair 10). The small variation in the number of submetacentric and subtelocentric in different populations was probably due to a differential contraction of the long and short arms of the chromosomes (Thiriot-Quiévreux, 1984b).

More recently, silver-staining of NORs (Nucleolar Organiser Regions) has been employed in *O. edulis* for a finer characterisation of chromosomal morphology. NORs correspond to sites of ribosomal DNA genes. A variable number of 1 to 4 Ag-NOR chromosomes was found, located terminally on the long arm of pairs 9 and 10 (Thiriot-Quiévreux and Insua, 1992). This study highlighted a strong heteromorphism in the apparent number of NORs per cell, the two most common configurations being one Ag-stained chromosome for both pairs 9 and 10, and 2 homologous Ag-stained chromosomes for the pair 10.

To our knowledge, no banding technique (C-banding or G-banding) has been employed in this species for a characterisation of heterochromatin patterns.

Finally, Rodríguez-Juíz et al. (1996) estimated by flow cytometry the diploid DNA content of *O. edulis* to be 2.33 ± 0.02 pg.

III- MATERIAL AND METHODS

III.1. Production of segregating families of *O. edulis* (Ifremer, La Tremblade, France)

The establishment of a genetic linkage map relies on the relationship between recombination frequency and genetic distance between markers and requires us to distinguish between parental and recombinant associations of these markers. The production of segregating families, with known parental genotypes, is therefore critical to the success of a genetic mapping project (See Chapter 1 for more details).

III.1.1. Experimental methodology for the achievement of biparental crosses in O. edulis

Because flat oysters are alternating hermaphrodites and females brood their larvae in their mantle cavity, it was not possible to determine in advance the sex of the oysters or to strip spawn the gonads for collecting gametes (contrary to classical methods of biparental crossing for the cupped oyster *C. gigas*). Therefore, the achievement of biparental crosses in *O. edulis* necessitates putting pairs of oysters in small aquariums, each aquarium being individually supplied with sea water. Moreover, to avoid the collection of larvae that do not come from the fertilisation of the pair of oysters placed into the aquarium (but from a fertilisation that occurred before the start of the experiment), each pair of oysters was kept for two weeks without a system of collection of larvae. Two weeks corresponds to the maximal brooding period indicated in the literature. After these two weeks, the outer water pipe of each aquarium was placed above a 100 µm-mesh sieve for the collection of larvae. The release of larvae into each sieve was checked daily and was easily recognised because larvae exhibit a characteristic slate-colour due to their shell. Each oyster was labelled; each aquarium exhibited the names of the two oysters placed into it and each sieve was labelled according to the aquarium (Figure 40).

Pairs of oysters that gave rise to the collection of larvae were killed and dissected for gill-samples that were preserved in 70% ethanol (diluted in water) for further genetic analysis.

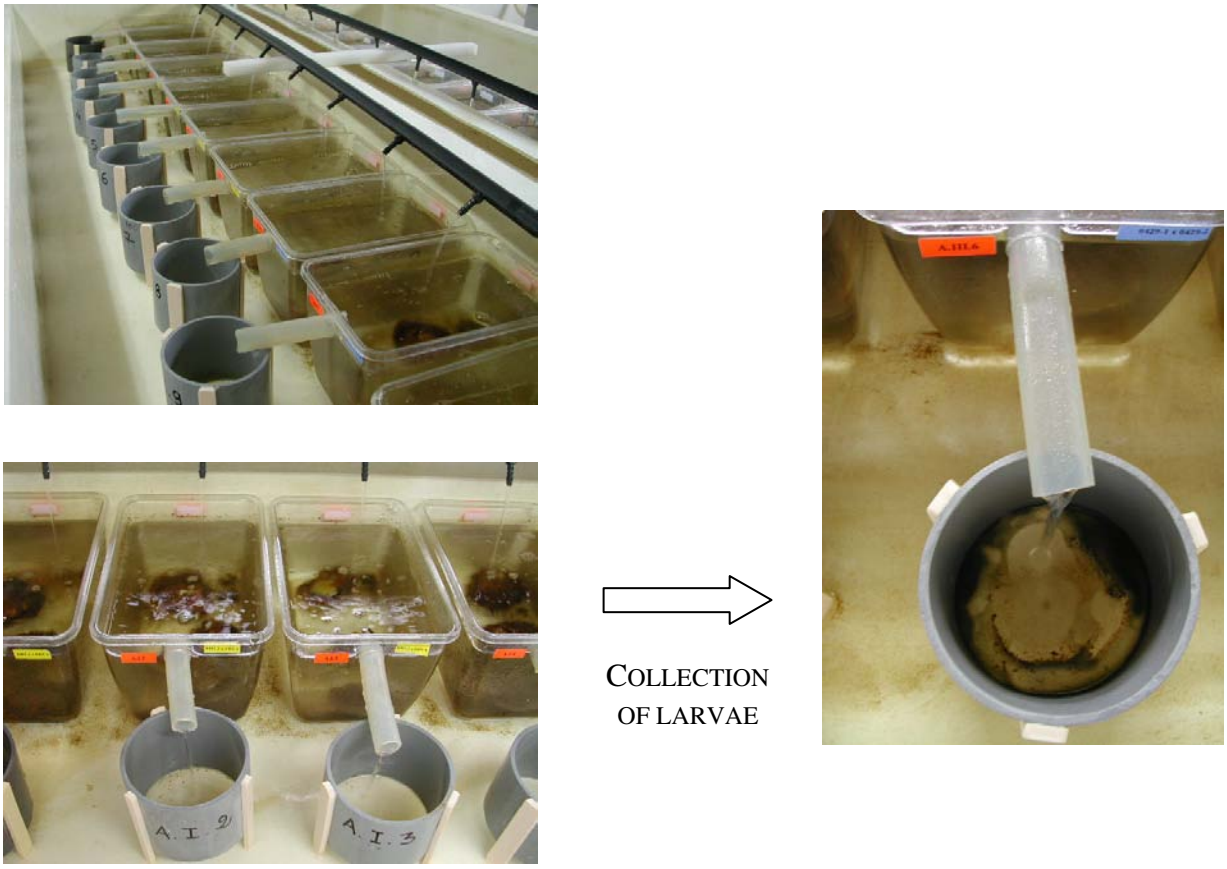


Figure 40. Experimental design for the production of biparental families in *O. edulis* (Ifremer, La Tremblade, France). Oysters were placed in small aquariums by pairs. Each aquarium had an individual supply of sea water and was labelled with the names of oysters and the name of the aquarium. The outer water pipe of each aquarium was placed above a 100 µm-mesh sieve for the collection of larvae (slate-coloured as shown on the picture on the right).

Each spawning (larvae collected in a sieve) was dated and transferred into 50 l cylinder tanks for the larval rearing that lasted around two weeks, at a concentration of around 150,000 larvae per tank. Water temperature was kept between 20.5 and 22°C. Phytoplankton (T-*Isochrysis* and *Chaetoceros gracilis*) was supplied daily. Food ration varied during the larval rearing and consisted of 20 cells.µl^{-1} from day (D) 1 to D7, 80 cells.µl^{-1} from D8 to D14 and 200 cells.µl^{-1} from D14 to D21. Water was changed three times a week. Due to growth differential of larvae belonging to the same spawning, up to three sieves of different mesh were superposed for collecting larvae during the water changing. The number of larvae in each spawning was counted three times a week and the overall growth of each spawning was checked by increasing the size of the sieve mesh during the period of larval rearing.

When more than 50% of ready-to-metamorphose larvae (size around 350 µm, exhibiting a mobile foot) were observed in a spawning, they were transferred in micronursery in 150 µm-mesh rectangular sieves (45x35x12 cm) placed into 150 l raceways. 120 ml of fine sterile crushed oyster shell (sieved on 200 to 300 µm and autoclaved) was placed into each 150 µm-mesh sieve to allow settlement of pediveliger larvae by providing them with a suitable substratum for their fixation. Water flow was set up at 150-200 $l.h^{-1}$ and water temperature at around 19-22°C. Phytoplankton was directly mixed upstream with the sea water and consisted of a mixture of 24 $cells.\mu l^{-1}$ of *Skeletonema costatum*, 2.7 $cells.\mu l^{-1}$ of T-*Isochrysis*, 1.6 $cells.\mu l^{-1}$ of *Chaetoceros gracilis* and 0.5 $cells.\mu l^{-1}$ of *Tetraselmis suecica* for 22-23 hours per day. Sieve mesh sizes were progressively increased according to the growth of oysters, from 150 µm to 1 mm (Figure 41).

For the purpose of genetic mapping of markers and QTLs mapping in the European flat oyster, two batches of production were achieved as described above at Ifremer, La Tremblade, France: families F2-L produced in 2004 (see below) and families F2-S in 2005 (detailed in Chapter 5). As my PhD project began in April 2004, the first year of production (2004) was performed by the technical team of Ifremer, La Tremblade, and in particular by Serge Heurtebise. I directly took part of the 2005 year of production, under the supervision of Serge Heurtebise.

By convention, the three-generations of the families produced will be noted as followed:

- F0 for the grand-parents, or F0 parents
- F1 for the parents, or F1 parents
- F2 for the progeny

III.1.2. Production of families F2-L (L=LINE)

In parallel to the programme of selection, 4 to 6 generations of inbred lines have been produced since 1995, by biparental crosses between fullsibs. Several inbred lines were produced, all initially derived from a cross between a selected oyster strain (S89ni, details in Chapter 5) and a wild oyster. Of particular interest, one inbred line, Oe 1-5-2-3-1, showed no mortality in the field since 2000 and was renamed OELL2000-set2 (Figure 42).

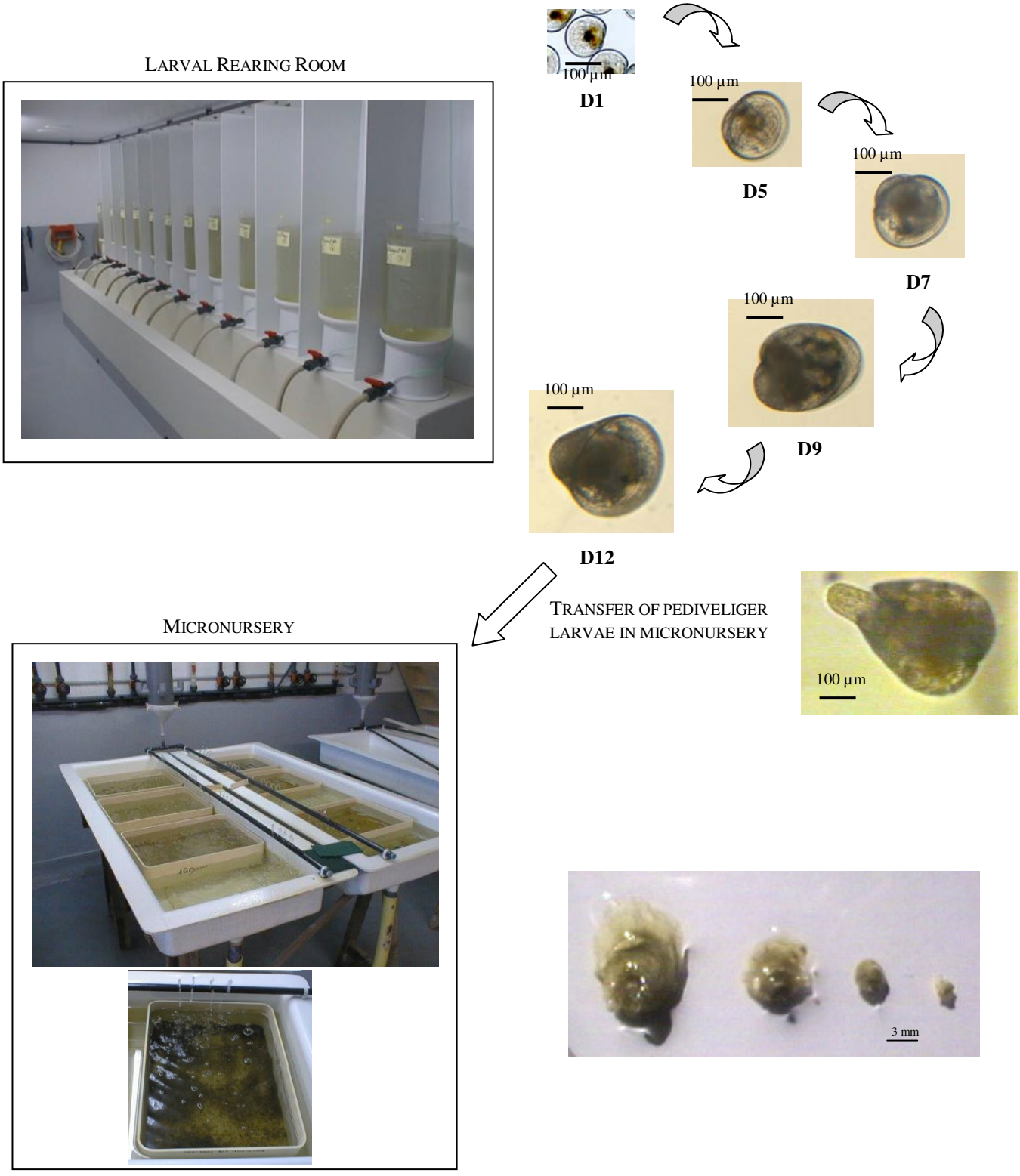


Figure 41. Larval rearing and micronursery of biparental families of *O. edulis* produced at the experimental hatchery of Ifremer, La Tremblade (France). D: number of days after release of larvae from the mantle cavity.

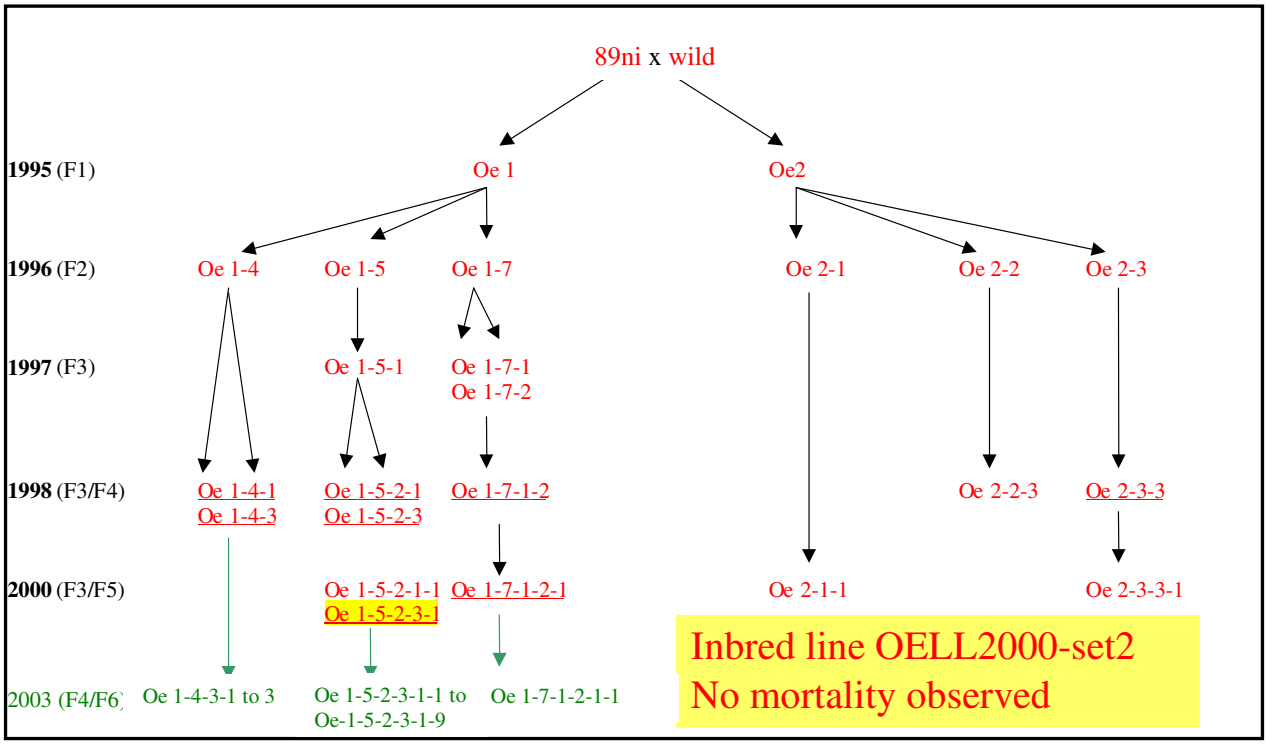


Figure 42. Inbred lines produced from the selected strain 89ni in *O. edulis*, by biparental crosses between fullsibs. One inbred line showed no mortality in the field or in hatchery since 2000, Oe 1-5-2-3-1, renamed OELL2000-set2.

In 2003, from 18 pairs of oysters (placed in individual aquariums as shown Figure 40), 4 families F1-L (L=LINE) were produced by crossing a wild oyster and an oyster from the inbred line OELL2000-set 2 that showed no mortality in the field (see Figure 42 for the full history of this inbred line). These four F1-L families were named OE.WL.03.21, OE.WL.03.23, OE.WL.03.27 and OE.WL.03.30; "OE" corresponding to the species *O. edulis*, "WL" to a cross between a wild oyster (W) and an oyster from a line (L), "03" to the year of production (2003) and the last two numbers to the batch. The same type of code for naming families will be applied later. In 2004, the second-generation F2-L has been obtained by fullsib crosses within each of the 4 F1-L families. Thirty pairs of oysters F1-L x F1-L were placed in individual aquariums, 7 for families OE.WL.03.21 and OE.WL.03.27, and 8 for families OE.WL.03.23 and OE.WL.03.30. Fourteen F2-L families were produced by the team of Ifremer, La Tremblade, named OE.F2.04.34, OE.F2.04.35, OE.F2.04.36, OE.F2.04.37, OE.F2.04.38, OE.F2.04.39, OE.F2.04.40, OE.F2.04.45, OE.F2.04.46, OE.F2.04.47, OE.F2.04.48, OE.F2.04.54, OE.F2.04.55 and OE.F2.04.63 (Figure 43).

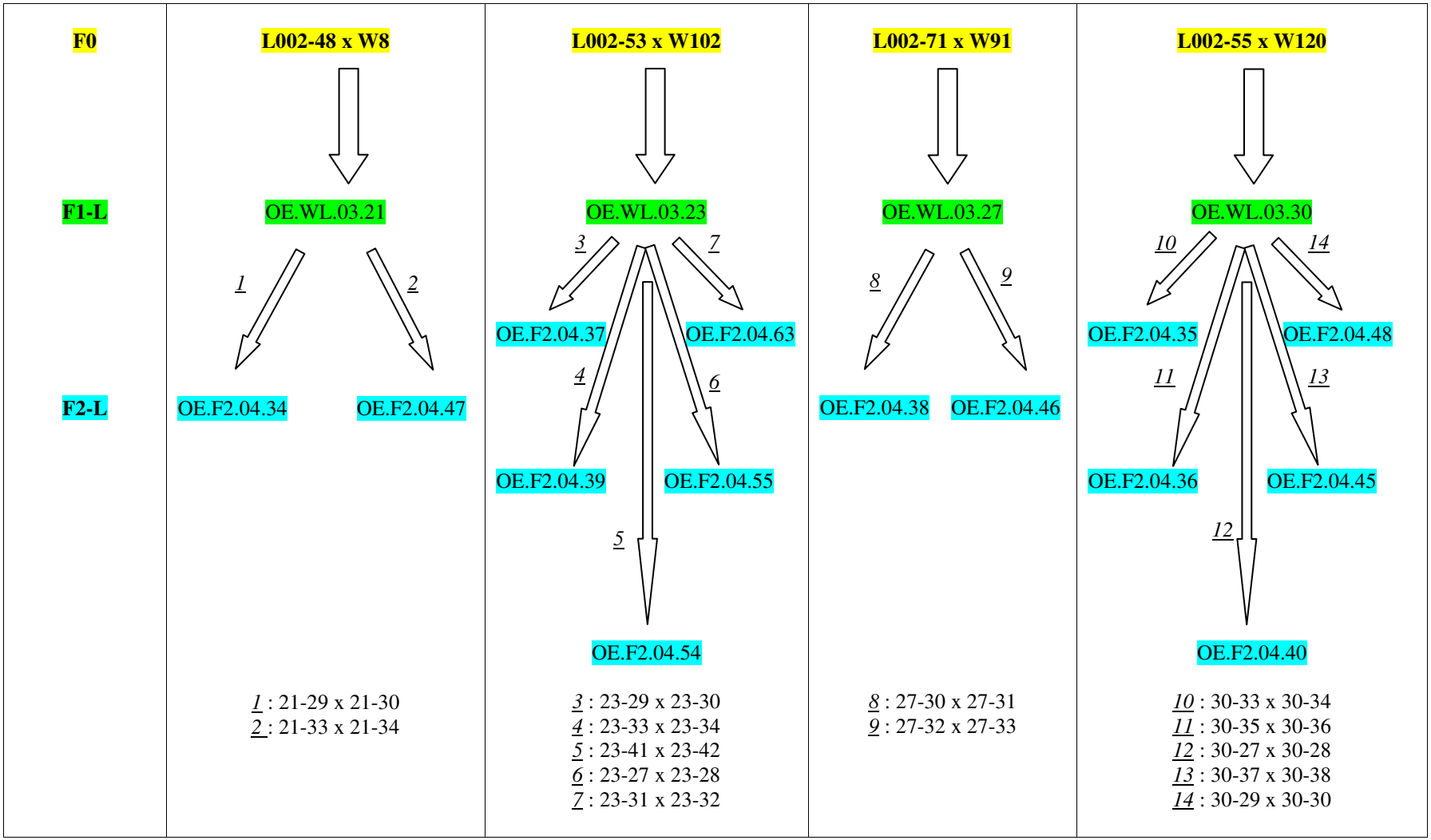


Figure 43. Production in 2004 of 14 F2-L families of *O. edulis* issued from crosses between an inbred line (OELL2000, see Figure 42) and wild oysters (W).

III.2. Choice of the mapping family

Fourteen F2-L families were produced in 2004 by the team of Ifremer, La Tremblade, France (see Figure 43 for more details). The aim of the study was to produce a first genetic linkage map in *O. edulis* and eventually to search for quantitative trait loci (QTL) of resistance to bonamiasis. In the context of QTL mapping, the traditional experimental design consists of crossing two divergent lines for the trait of interest, for example between a high-selected line and a low-selected line. As no "lines" are available in *O. edulis*, the first-generation production (F1-L) was obtained by crossing 2 individuals supposedly divergent, one individual issued from an inbred line that showed no mortality in the field (supposedly resistant to bonamiasis) and an individual from the wild (whose resistance is unknown).

Molecular markers available in the flat oyster *O. edulis* consist of 20 microsatellites and 60 different AFLP primer combinations (see Chapter 2 for protocols and Table 12).

The first step of the genetic and QTL mapping project required the choice of an F2-L family from the 14 produced, based mostly on DNA polymorphism and amount of DNA extracted.

Therefore, 20 microsatellites and 60 AFLP primer pairs were first genotyped for the 8 F0 parents. Marker polymorphism was assessed based on the F0-parents genotypes. For a given pair of parents, microsatellite markers were classified as:

- fully informative (FI) when there was no common allele between the two parents and that both parents were heterozygous for different alleles (for example 106/112 for the first parent and 102/108 for the second parent) (Figure 44)
- semi-informative (SI) when there was one allele in common (for example 142/148 for the first parent and 142/158 for the second parent); or when one parent was heterozygous for two alleles and the other parent was homozygous for a different allele
- non informative (NI) when both parents exhibited the same genotype.

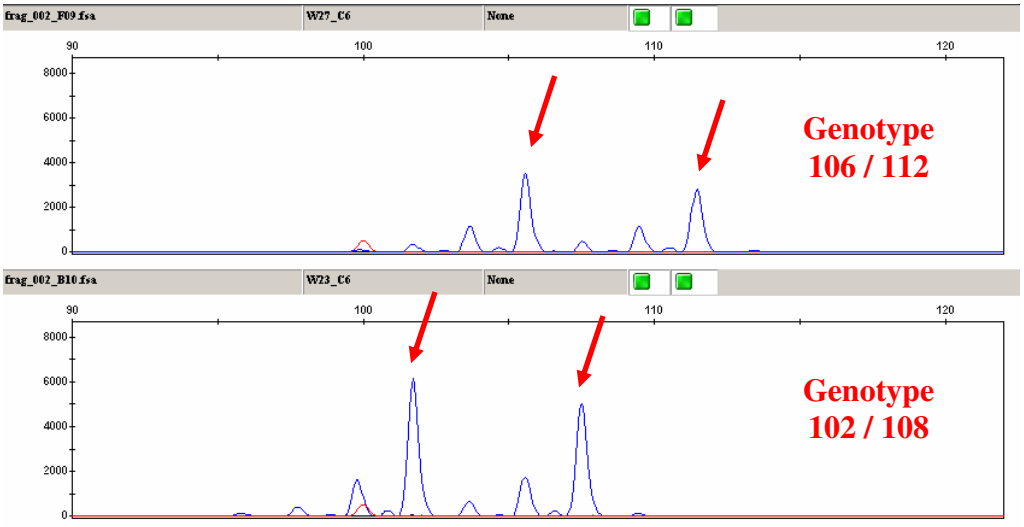


Figure 44. Example of a fully informative microsatellite marker (*Oedu*C6) for a pair of parents.

For each pair of parents, the number of polymorphic AFLP markers was counted. A polymorphic site was identified when the peak was present for one of the two parents and absent in the other parent (Figure 45). The most polymorphic F1-L family was the one whose parents exhibited the highest number of fully informative microsatellites and the maximum number of polymorphic AFLPs.

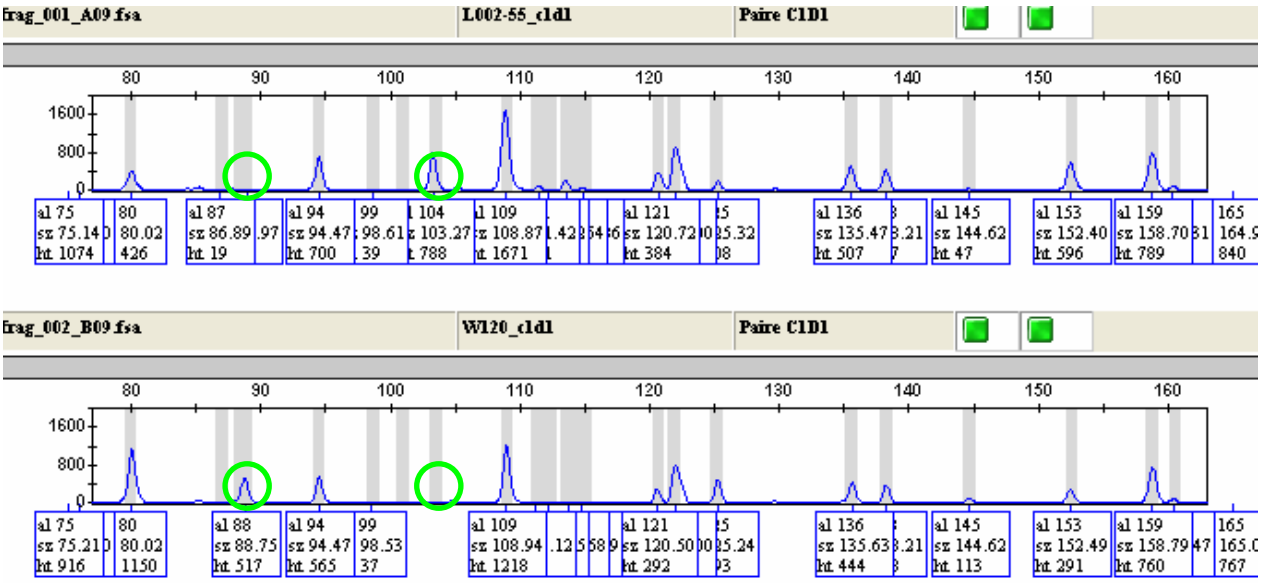


Figure 45. Identification of the polymorphic AFLP markers for a pair of parent, L002-55 and W120. Two polymorphic markers are shown, allele 88 and allele 104.

Then, the 20 microsatellites were scored on the F1-parents of the F1-L family chosen. The F2-L family was chosen according to microsatellite polymorphism of the F1 parents, based on three criteria:

- Criterion 1: for the fully informative loci, select the F2 family issued from F1 parents that keep the maximum number of alleles of the parents (F0).
- Criterion 2: for the semi informative loci, try to select F2 family where the common allele is not present in the F1 parents to avoid ambiguity of origin during segregation analysis in the F2.
- Criterion 3: for loci which exhibit a null allele, select F1 parents for which genotype is not ambiguous due to the segregation of this null allele.

III.3. Segregation distortion analysis

Segregation distortion analysis was achieved by a χ^2 goodness of fit statistical test, which relies on the following computation:

$$\chi^2 = \sum_{i=1}^{n} \frac{(Obs - Exp)^2}{Exp}$$

where *i* is the iterative number, *n* the number of classes, *Obs* the observed number and *Exp* the expected number in the class. Depending on the type of marker (codominant or dominant), and the type of segregation, the computation of the χ^2 value will be summed on a variable number of classes (from *i*=2 to 4, see below). Segregation distortion was estimated between the F1 parents and the F2 progeny.

For microsatellites, four types of segregation could be observed depending on the number of alleles present in the two parents:

- 1:1:1:1 type: correspond to FI microsatellites, where 4 alleles are present (both parents being heterozygous for 2 different alleles) and to SI microsatellites, where 3 alleles are present (both parents being heterozygous, with one allele in common)
- 1:2:1 type: correspond to microsatellites for which both parents are heterozygous for the same 2 alleles

- 1:1 type: correspond to microsatellites exhibiting 3 alleles between the two parents, one parent being heterozygous for the first two alleles and the second parent being homozygous for the third allele and to microsatellites exhibiting 2 alleles between the two parents (Table 21).

Table 21. Different segregation types of microsatellites observed between the F1 parents and their F2 progeny. a1, a2, a3, a4 are the 4 possible alleles at a given microsatellite locus, FI: fully informative, SI: semi informative, NI: non informative, N: total number of offspring genotyped, ν: degree of freedom (number of classes-1), n_{obs1}, n_{obs2}, n_{obs3} correspond to the observed number of progeny exhibiting the genotype and n_{exp1}, n_{exp2}, and n_{exp3} correspond to the expected number of progeny in each genotypic class.

Segregation type	Number of alleles	Parents genotypes	Offspring genotypes (Observed classes)	Expected number of offspring genotypes	ν
1:1:1:1 FI	4	a1a2 x a3a4	a1a3 a1a4 a2a3 a2a4	N/4 N/4 N/4 N/4	3
1:1:1:1 SI	3	a1a2 x a1a3	a1a1 a1a3 a1a2 a2a3	N/4 N/4 N/4 N/4	3
1:2:1 NI	2	a1a2 x a1a2	a1a1 (n_{obs1}) a1a2 (n_{obs2}) a2a2 (n_{obs3})	N/4 (n_{exp1}) N/2 (n_{exp2}) N/4 (n_{exp3})	2
1:1 SI	3	a1a2 x a3a3	a1a3 a2a3	N/2 N/2	1
1:1	2	a1a2 x a1a1	a1a1 a1a2	N/2 N/2	1

For example, the χ^2 value for a 1:2:1 type microsatellite was computed as:

$$\chi^2 = \frac{(n_{obs1} - n_{exp1})^2}{n_{exp1}} + \frac{(n_{obs2} - n_{exp2})^2}{n_{exp2}} + \frac{(n_{obs3} - n_{exp3})^2}{n_{exp3}}$$

Depending on the degree of freedom, decision rules were applied as stated in Table 22.

Table 22. Decision rules for the χ^2 goodness of fit statistical test, depending on the degree of freedom (ν). NS: test not significant; p<0.05, p<0.01 and p<0.001 correspond to significant tests at a probability α=5%, 1% or 1‰.

ν=1	ν=2	ν=3	Test	Interpretation
$\chi^2 < 3.841$	$\chi^2 < 5.991$	$\chi^2 < 7.815$	NS	Mendelian marker
$3.841 < \chi^2 < 6.635$	$5.991 < \chi^2 < 9.210$	$7.815 < \chi^2 < 11.345$	p<0.05	Distorted marker
$6.635 < \chi^2 < 10.828$	$9.210 < \chi^2 < 13.816$	$11.345 < \chi^2 < 16.266$	p<0.01	Highly distorted marker
$\chi^2 > 10.828$	$\chi^2 > 13.816$	$\chi^2 > 16.266$	p<0.001	Very highly distorted marker

For the AFLP markers, which are dominant, there are only two genotypic classes, presence or absence of the peak. Two types of segregation can be observed, either 1:1 (when only one of the two parents exhibits the peak) or 3:1 (when both parents exhibit the peak). A χ^2 goodness of fit statistical test was applied to the 1:1 and 3:1 markers to detect segregation distortion:

$$\chi^2 = \frac{\left(O_{peak} - E_{peak}\right)^2}{E_{peak}} + \frac{\left(O_{nopeak} - E_{nopeak}\right)^2}{E_{nopeak}}$$

where O_{peak} and O_{nopeak} are the observed numbers of individuals with or without peak; and E_{peak} and E_{nopeak} the expected number of individuals with or without peak.

In the case of type 1:1 markers, $E_{peak}=E_{nopeak}=N/2$ with N the number of F2 progeny. In the case of type 3:1 markers, $E_{peak}=3N/4$ and $E_{nopeak}=N/4$. This test follows a χ^2 distribution with one degree of freedom ($\nu=1$). See Table 22 for $\nu=1$ for the decision rule applied.

As the F2-L families produced were of three-generation pedigree type (grand-parent or F0 parent, F1 parents and F2 progeny), inference of AFLP genotypes from the phenotype (presence or absence of peak) allowed the distinction of several configurations of transmission of AFLPs:

- type I for which there were no missing data (each phenotype could be assigned to a genotype, *Aa* or *aa*) and the grand-parental and parental origins of AFLP alleles could be tracked without ambiguity (Figure 46)
- type II for which there were no missing data (each phenotype could be assigned to a genotype, *Aa* or *aa*) but there was ambiguity for tracking the grand-parental origin of the *A* allele because both F0 parents were *Aa* (Figure 46)
- type III for which the presence of a peak lead to an ambiguous genotype *A?* (either *AA* or *Aa*) and the grand-parental origin of the *A* allele could be tracked because only one of the two F0 parents exhibited this allele (Figure 47)
- type IV for which the presence of the peak was associated with the ambiguous genotype *A?* but the grand-parental and parental origins could not be tracked because all four F0 and F1 parents exhibited the peak (Figure 47). This configuration was considered not informative and AFLPs of this type were discarded from further analysis.

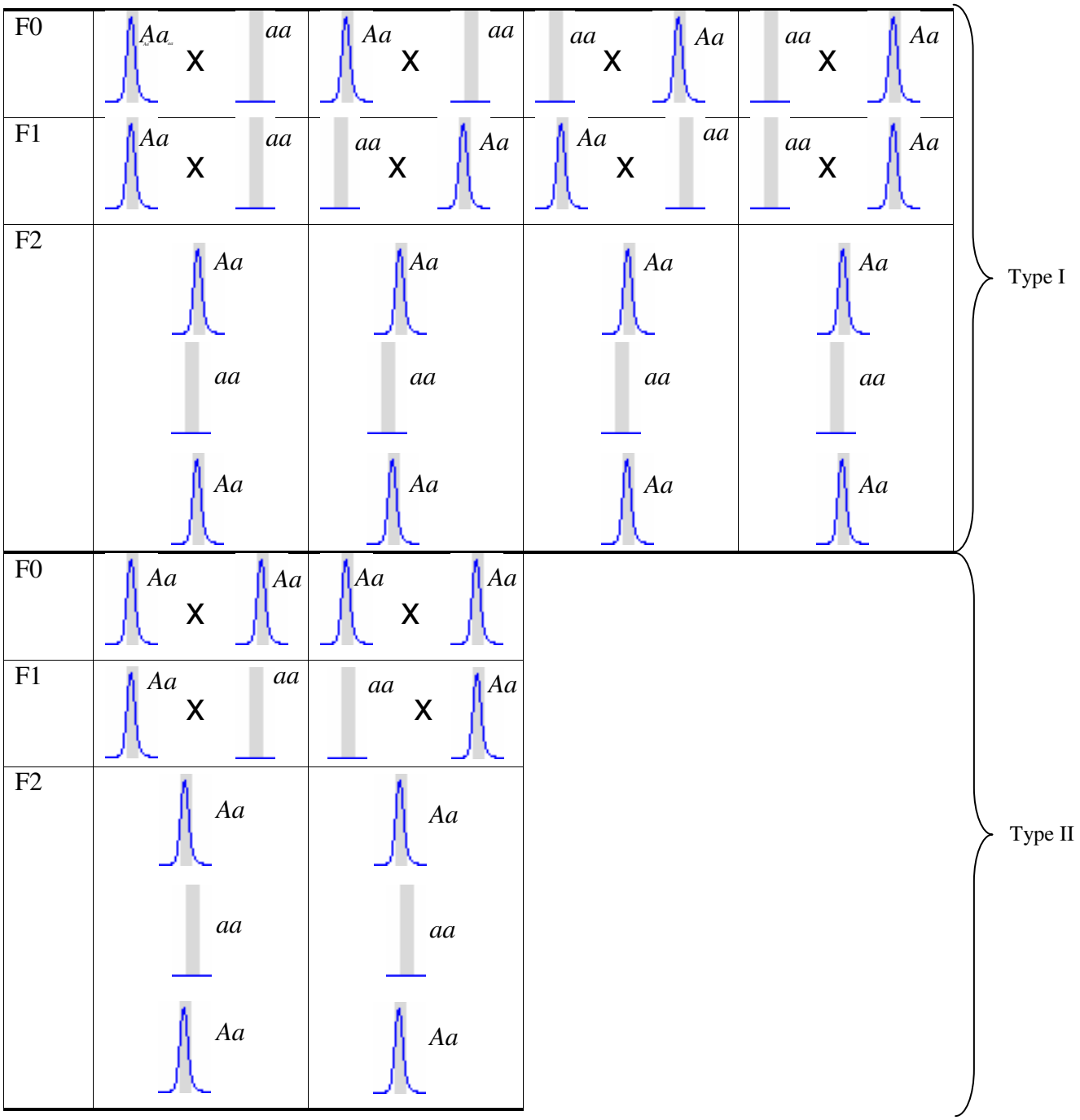


Figure 46. Three-generation configurations of segregation for the 1:1 AFLPs. *Aa* corresponds to the genotype associated with the presence of the peak and *aa* to the genotype associated to the absence of the peak. In F0 and F1 are represented the genotypes of the two parents generating the next generation (first parent on the left, second parent on the right, X symbolising the cross between the two parents). In F2, 3 F2 progeny are shown, 2 of them exhibiting the peak, the third one not exhibiting the peak, to illustrate that only AFLPs that are segregating in the F2 progeny are considered.

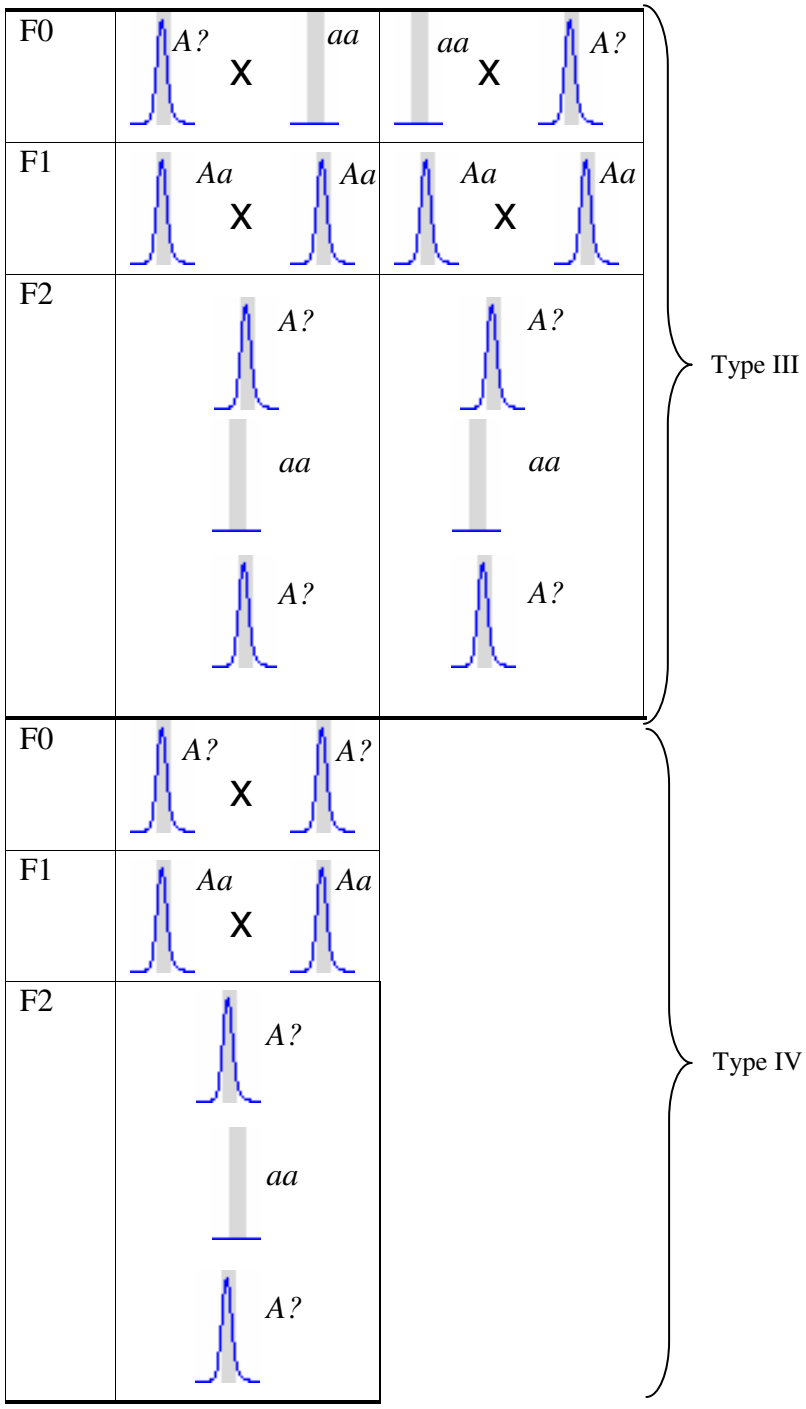


Figure 47. Three-generation configurations of segregation for the 3:1 AFLPs. See Figure 46 for the main explanations. *A?* corresponds to an ambiguous genotype associated with the presence of the peak (either *AA* or *Aa*).

III.4. CRIMAP software

CriMap software (Green et al., 1990) allows the construction of multilocus linkage maps, the assessment of support relative to alternative locus orders by generating LOD score tables and the detection of probable data errors. This software has been run under a UNIX environment. I received training at the Roslin Institute (Edinburgh, Scotland, UK) under the supervision of Dr. Chris Haley, thanks to the award of a one-month fellowship (Gender Action Plan, Marine Genomics).

CriMap software requires only one input file, "chr*.gen", which contains the raw genotype data. The structure of the file includes the number of families, the number of markers, the name of the markers and then the genotypic data for each individual of each family. This software is very flexible on the way genotypic data can be entered and does not require a specific coding for the data or a particular structure of pedigree (in contrast to MapMaker software, see Chapter 3). Therefore, instead of using random numbers associated with each allele (1, 2 …), we used the length of the fragment in base pair. For microsatellites, the two alleles were entered as the fragment lengths (for example 144 148 for a heterozygous genotype). For the AFLPs, a genotype *Aa* was noted "253 1" for an AFLP of length 253 bp, the "1" referring to the null allele; a genotype *aa* was noted "1 1" and a genotype *A?* was noted "253 0", "0" being a missing value. An individual for which the PCR amplification failed at a particular locus was noted "0 0". An example of input file is shown Figure 48.

Then the PREPARE command allows the creation of 4 files used by the software internally: "chr*.par" contains the parameter for the analysis (e.g. sex-averaged or sex-specific analysis, sequence of an ordered set of markers, name of loci to be inserted into the map) and can be modified by the user; "chr*.loc" contains the marker informativity (score of the number of informative meioses); "chr*.dat" contains the restructured datafile and "chr*.ord" contains the marker orders. These files are required by the programme to perform the commands described below.

```
1                    "Number of families analysed"
9                    "Number of markers"
J12   U2    HA1   C6    O9    A12-288     B1-310      B1-339      B2-186      "Names of markers"

OE.F2.04.63          "Name of the family"
8                    "Number of individuals in the family"

1001  0     0     0     "Parent F0: 1001" "Its two parents are unknown: 0" "Its sex is unknown: 0"
248   258   156   176   154   154   108   120   156   164   1     1     310   1     1     1     186   0

1002  0     0     0     "Parent F0: 1002" "Its two parents are unknown: 0" "Its sex is unknown: 0"
232   242   156   204   132   144   102   106   162   168   288   1     310   1     339   1     1     1

1003  1001  1002  0     "Parent F1: 1003" "Its two parents are 1001 and 1002" "Its sex is unknown: 0"
232   258   176   204   132   154   106   108   164   168   288   1     310   1     339   1     186   1

1004  1001  1002  0     "Parent F1: 1004" "Its two parents are 1001 and 1002" "Its sex is unknown: 0"
242   258   156   176   144   154   102   108   156   162   1     1     1     1     1     1     186   1

9     1003  1004  0     "Progeny F2: 9" "Its two parents are 1003 and 1004" "Its sex is unknown: 0"
232   258   176   204   132   154   106   108   156   168   288   1     310   1     339   1     186   0

12    1003  1004  0     "Progeny F2: 12" "Its two parents are 1003 and 1004" "Its sex is unknown: 0"
242   258   176   176   154   154   108   108   162   164   1     1     310   1     1     1     186   0

13    1003  1004  0     "Progeny F2: 13" "Its two parents are 1003 and 1004" "Its sex is unknown: 0"
232   258   176   204   132   154   106   108   156   168   288   1     1     1     339   1     186   0

14    1003  1004  0     "Progeny F2: 14" "Its two parents are 1003 and 1004" "Its sex is unknown: 0"
242   258   156   204   144   154   102   106   162   164   288   1     310   1     1     1     186   0
```

Figure 48. Example of input file for CriMap software, for 9 molecular markers (5 microsatellites and 4 AFLPs) for 8 individuals, in a three-generation pedigree (2 F0 parents named 1001 and 1002, 2 F1 parents named 1003 and 1004 and 4 F2 progeny named 9, 12, 13, 14). In turquoise are shown the microsatellites genotypes (e.g. the F0 parent 1001 is heterozygote 248 258 for J12). In red are shown the genotypes at two AFLP markers, A12-288 and B1-339 (type I). In grey are shown the genotypes at one AFLP marker, B1-310 (type II) and in dark blue at a type III AFLP marker (B2-186). Type I, type II and type III refer to Figures 46 and 47. The explanations are put in brackets and in bold. "0" represents missing genotype, "1" the null allele *a* of an AFLP. Markers are named according to the length of the fragment amplified.

The first step of the linkage analysis consisted of the TWO POINT command which estimated recombination frequencies between each pair of markers for each of the two F1 parents, at a LOD score ≥ 3.0. These two-point data were used to construct sex-specific linkage groups.

Once linkage groups have been found, further analysis was done separately for each of the two F1 parents, and group by group. It means that a new input file "chr*.gen" had to be done for each linkage group of each F1 parent, each file containing the genotypic data of the markers present in the linkage group under study.

The second step of linkage analysis consisted of ordering markers within each linkage and several commands were used for this purpose. Firstly, the BUILD command was performed and lead to the construction of a map for each linkage group by sequential incorporation of loci, starting with the most informative pair of markers. Then the FIXED and ALL commands were achieved in order to add the unplaced markers after BUILD to the map, by decreasing order of informativeness. By analogy with MapMaker, BUILD is the equivalent to the THREE POINT command, FIXED and ALL to the TRY command. After the addition of a new marker to a sequence of ordered markers, the new order was tested against alternative order by the FLIPS command (RIPPLE in MapMaker). Markers that lead to ambiguous map position were discarded from further analysis. Finally, any suspect markers could be detected with the CHROMPIC command which displayed the number and location of recombinations on each chromosome, highlighting candidate data errors.

III.5. Genome length and map coverage

Average marker spacing of each map was calculated by dividing the total length of the map by the number of intervals (the number of markers minus the number of linkage groups). In the same way, the average marker spacing for each linkage group was calculated by dividing the length of each linkage group by the number of intervals on that linkage group (i.e. the number of markers minus 1).

The expected length of the genome was estimated using method 4 of Chakravarti et al. (1991): the length of each linkage group was multiplied by the factor (m+1)/(m-1); m being the number of markers on each linkage group. Observed genome coverage estimates were determined by dividing the observed genome length by the expected length of the genome.

III.6. Recombination frequency differences

Differences in recombination frequencies between the two parents were estimated using G-tests of independence that compared parental and recombinant genotypes for each parent for each pair of linked markers (with LOD score > 3.0).

IV- RESULTS

IV.1. Choice of the mapping family

Firstly, choice was made based on the DNA concentration and microsatellite polymorphism in the F0 parents. DNA extraction from three individuals, L002-48, W8 and L002-71, was insufficient to achieve microsatellite and AFLPs amplifications. This was due to a very small amount of tissue available for this DNA extraction (these individuals have been biopsed alive, under $MgCl_2$ anesthesia). Two couples exhibited 14 and 15 fully informative microsatellites on 22 amplified: L002-53 and W102 (parents of the F1-L family OE.WL.03.23); L002-55 and W120 (parents of the F1-L family OE.WL.03.30) (Table 23). Microsatellite amplifications also allowed the confirmation of parentage between the F0 and F1 parents (Figure 49).

Secondly, the number of polymorphic AFLP markers was counted between the two pairs of F0 parents of the 2 F1-L families chosen on microsatellite polymorphism, OE.WL.03.23 and OE.WL.03.30. Parents L002-53 and W102 of OE.WL.03.23 exhibited 642 polymorphic sites while parents L002-55 and W120 of OE.WL.03.30 exhibited 725 polymorphic sites (Table 24). Therefore, based on microsatellite and AFLP polymorphisms, the F1-L family OE.WL.03.30 was chosen.

Table 23. Choice of an informative F1-L family based on microsatellite polymorphism of the F0 parents.
C_{DNA}: DNA concentration. N_a: number of microsatellites amplified. FI: fully informative (no common allele between the two parents). SI: semi-informative (one allele in common). NI: non informative (two alleles in common). Nb: number.

F1 Family name	Parent name	C_{DNA} µg/mL	N_a	Loci FI	Loci SI	Loci NI	Nb loci FI	Nb loci SI	Nb loci NI
OE.WL.03.21	L002-48	70	13	B0, J12, U2, H15, G9	C6, O9, T5, HA7, Oe3/37	-	5	5	-
	W8	100	12						
OE.WL.03.23	L002-53	235	22	B11, C6, J12, T5, O9, G9, T18, HA1, HA7, HA21, HA11, Oe1/47, Oe1/63, Oe3/44	U2, H15, Oe2/71, Oe3/37	Oe1/10 Oe1/64	14	4	2
	W102	408	22						
OE.WL.03.27	L002-71	38	12	B11, J12, O9, G9, HA7, Oe1/47	B0, C6, U2, H15, T5	Oe3/37	6	5	1
	W91	237	21						
OE.WL.03.30	L002-55	262	22	B11, C6, J12, U2, H15, T5, O9, G9, HA1, HA7, HA21, HA11, Oe1/47, Oe1/63, Oe1/64	T18, Oe2/71, Oe3/44	B0 Oe1/10 Oe3/37	15	3	3
	W120	327	22						

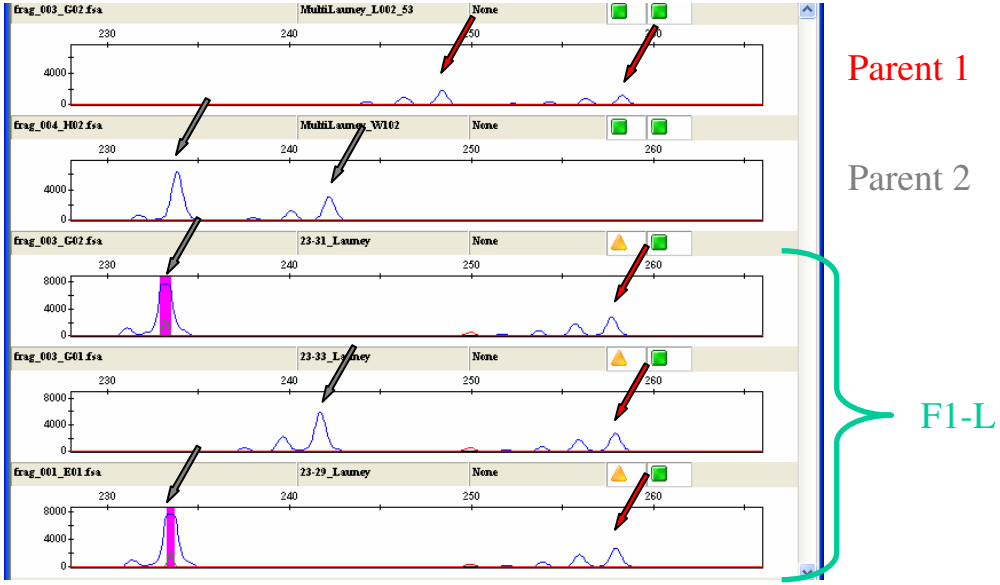


Figure 49. Confirmation of parentage between the F0 and F1 parents of families F2-L based on microsatellite genotyping. Example for the family OE.WL.03.23 at *Oedu*J12 microsatellite locus. The pink coloration corresponds to a very high intensity signal.

Table 24. Choice of an F1-L family based on AFLPs polymorphism of the F0 parents. Sixty AFLP primer pairs were genotyped for the two pairs of F0 parents chosen based on microsatellite polymorphism. See Table 12 for the abbreviations of the primer pairs used.

Family name Primer pair	OE.WL.03.23 (L002-53xW102)	OE.WL.03.30 (L002-55xW120)	Family name Primer pair	OE.WL.03.23 (L002-53xW102)	OE.WL.03.30 (L002-55xW120)
A1	13	16	**D1**	19	10
A2	8	10	**D2**	9	11
A3	4	10	**D3**	?	25
A4	11	11	**D4**	11	9
A5	7	8	**D5**	7	15
A6	8	?	**D6**	19	24
A7	19	17	**D7**	22	12
A8	4	7	**D8**	11	13
A9	14	5	**D9**	11	19
A10	12	18	**D10**	16	18
A11	9	13	**D11**	7	?
A12	12	13	**D12**	12	11
B1	5	6	**E1**	16	19
B2	3	5	**E2**	4	14
B3	3	6	**E3**	4	8
B4	12	15	**E4**	15	9
B5	16	11	**E5**	6	12
B6	8	13	**E6**	14	14
B7	15	11	**E7**	19	22
B8	10	13	**E8**	12	7
B9	10	5	**E9**	6	10
B10	8	8	**E10**	11	13
B11	8	12	**E11**	8	9
B12	6	6	**E12**	8	11
C1	17	17	**TOTAL**	**642**	**725**
C2	16	12			
C3	12	13			
C4	13	17			
C5	14	16			
C6	24	17			
C7	13	19			
C8	3	13			
C9	16	17			
C10	9	14			
C11	13	?			
C12	?	16			

Finally, once the F1-L family has been chosen, an F2-L family was chosen as the mapping family based on microsatellite polymorphism in the F1 parents. Indeed, for the family OE.WL.03.30, 8 pairs of F1 parents were put into individual aquariums; 5 F2-L families were produced, derived from 5 different biparental crosses: OE.F2.04.35, OE.F2.04.36, OE.F2.04.40, OE.F2.04.45 and OE.F2.04.48 (see Figure 43).

According to the criterion 1, 2 F2-L families could be chosen, OE.F2.04.36 and OE.F2.04.45, because they exhibited a similar mean number of alleles per fully informative locus between the two F1 parents, 3.08 and 3.07 respectively. According to the criterion 2, 2 (of 3) loci that were semi informative in the F0 parents became fully informative in the F1 parents 30-35 and 30-36. The origin of each allele could be tracked; there would be no ambiguity of parental allele origin in the F2 generation. For the F1 parents 30-37 and 30-38, this was the case for only one (of 3) loci. So there was a little advantage at this step for the F2-L family OE.F2.04.36 derived from the cross between 30-35 and 30-36. For the criterion 3, the F1 parents 30-35 and 30-36 exhibited no ambiguous genotypes, even when a null allele was segregating. Nevertheless, the genotype was unknown for 30-36 at Oe1/64 locus (only 19 microsatellites amplified for this individual). On the contrary, individuals 30-37 and 30-38 (OE.F2.04.45) have both been amplified for 21 microsatellites; but there was one ambiguous genotype (for 30-37 at *Oedu*T18 locus) (Table 25).

Therefore, two F2-L families, OE.F2.04.36 and OE.F2.04.45, were potential candidates to be chosen as the mapping family, with a little advantage for OE.F2.04.36. However, at the end of September 2004, evaluation of the stocks produced in 2004 revealed that high mortality occurred in the family OE.F2.04.36 because only two bags of 231 and 257 oysters could be kept in an outdoor tank at the experimental hatchery of Ifremer, La Tremblade. In contrast, 1000 oysters were randomly picked up and put into 2 bags (500 oysters per bag) for the family OE.F2.04.45.

In conclusion, the F2-L family named OE.F2.04.45 was chosen as the mapping family for further genetic linkage analysis studies.

Table 25. Choice of the F2-L mapping family based on microsatellite genotypes for the F1 parents knowing the F0 parents' genotypes.
C_{DNA}: DNA concentration. N_a: number of microsatellites amplified. FI: fully informative. SI: semi informative. ?:missing data.
Criterion 1: mean number of alleles per fully informative locus between the two F1 parents. Criterion 2: number of locus semi informative in the F0 parents which are fully informative in the F1 parents (no more ambiguity of allele origin in the F2).
Criterion 3. Null: presence of a null allele, yes (Y) or no (N). Amb.: ambiguous genotype, yes (Y) or no (N).

F2 family name	Parent name	C_{DNA} µg/mL	N_a	Criterion 1: Loci FI	Criterion 2 : Loci SI	Criterion 3 : presence of null alleles									
						H15		T18		G9		B11		Oe1/64	
						Null	Amb.	Null	Amb.	Null	Amb.	Null	Amb.	Null	Amb.
OE.F2.04.35	30-33	746	21	2.50 (14/15 loci)	0	Y	N	Y	Y	Y	N	Y	N	Y	N
	30-34	612	20			Y	N	Y	Y	Y	N	N	N	Y	N
OE.F2.04.36	30-35	471	21	3.08 (13/15 loci)	2 of 3 (T18, Oe3/44)	Y	N	Y	N	N	N	N	N	Y	N
	30-36	768	19			Y	N	Y	N	N	N	Y	N	?	?
OE.F2.04.40	30-27	495	21	2.73 (15/15 loci)	1 of 3 (Oe3/44)	N	N	Y	Y	N	N	N	N	Y	N
	30-28	725	21			Y	N	Y	N	N	N	N	N	Y	N
OE.F2.04.45	30-37	671	21	3.07 (15/15 loci)	1 of 3 (Oe2/71)	Y	N	Y	Y	N	N	N	N	Y	N
	30-38	485	21			Y	N	N	N	N	N	Y	N	Y	N
OE.F2.04.48	30-29	269	20	2.64 (14/15 loci)	1 of 3 (Oe2/71)	N	N	Y	Y	Y	N	N	N	Y	N
	30-30	881	20			Y	N	?	?	Y	N	N	N	Y	N
OE.F2.04.37	23-29	529	21	2.55 (11/14 loci)	2 of 4 (U2, H15)	N	N	Y	N	N	N	Y	N	N	N
	23-30	490	18			N	N	?	?	N	N	Y	N	N	N
OE.F2.04.39	23-33	492	21	2.71 (14/14 loci)	1 of 4 (U2)	Y	Y	N	N	N	N	Y	N	N	N
	23-34	549	21			Y	Y	Y	N	N	N	Y	N	N	N
OE.F2.04.63	23-31	587	20	3.25 (12/14 loci)	1 of 4 (U2)	Y	Y	?	?	N	N	Y	N	N	N
	23-32	579	19			Y	Y	?	?	N	N	N	N	N	N

IV.2. Parentage analysis

Parentage analysis performed on 20 microsatellite markers revealed that family OE.F2.04.45 came in fact from a self-fertilisation of an F1 individual, 30-37, and not from a biparental cross (30-37 x 30-38) as thought. For example, at locus *Oedu*J12, the two F0 parents were 248/258 (L002-55) and 226/244 (W120); the two F1 parents were 244/248 (30-37) and 244/258 (30-38). Therefore, 4 equiprobable genotypes were expected in the F2 progeny: 244/244, 244/258, 244/248 and 248/258. Contrary to expectations, only three genotypes were found in the F2 progeny analysed (n=96): 244/248, 244/244 and 248/248 (Figure 50). Because the family OE.F2.04.45 was part of the "*Bonamia*-challenge experiment" (See Chapter 5), this family was not discarded from further linkage analysis. However, only 48 F2 individuals were scored for the AFLPs instead of 96 initially planned. In addition, another family was chosen for genetic mapping because of its level of polymorphism (Table 25) and the high number of oysters that survived: OE.F2.04.63 that came from a real biparental cross of 2 F1 individuals.

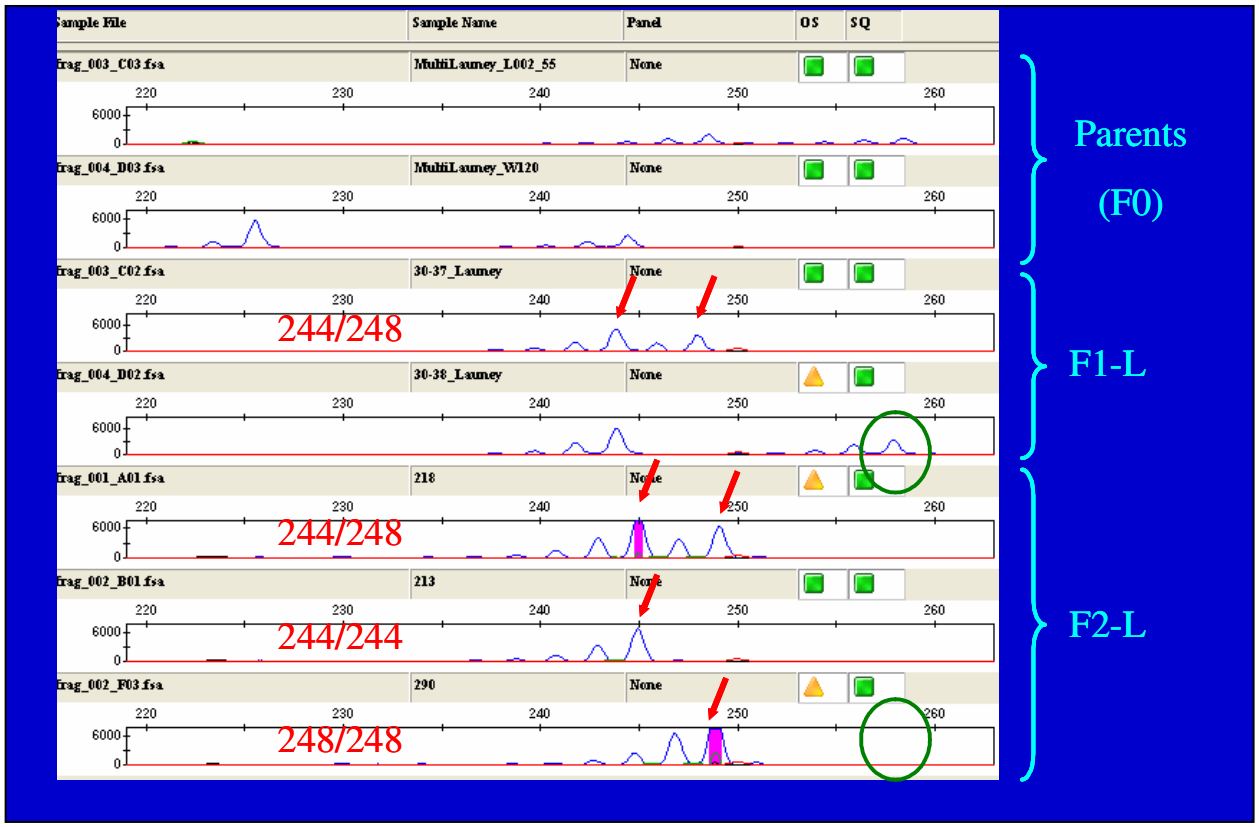


Figure 50. Evidence of self-fertilisation of the F1 individual 30-37 to produce the OE.F2.04.45 family: example at the *Oedu*J12 microsatellite. The two F0 parents were 248/258 (L002-55) and 226/244 (W120). The two F1 parents were 244/248 (30-37) and 244/258 (30-38). Contrary to expectations, only three genotypes were found in the F2 progeny analysed (n=96): 244/248, 244/244 and 248/248.

IV.3. Segregation distortion

For the first mapping family, OE.F2.04.45, 48 F2 progeny were genotyped for 20 microsatellites and 60 AFLP primer pairs. Very high segregation distortion was reported for this family because only 17% (3 out of 18 informative markers: *Oedu*T18, *Oedu*.HA11a and Oe1/64) of the microsatellites and 50% of the AFLP markers (213 out of 424 markers) were segregating according to Mendelian rules (Figure 51). All AFLPs were of 3:1 segregation type because this F2 family came from the self-fertilisation of one F1 parent. In total, 231 markers were kept for further genetic linkage analysis, 18 microsatellites and 213 Mendelian AFLPs.

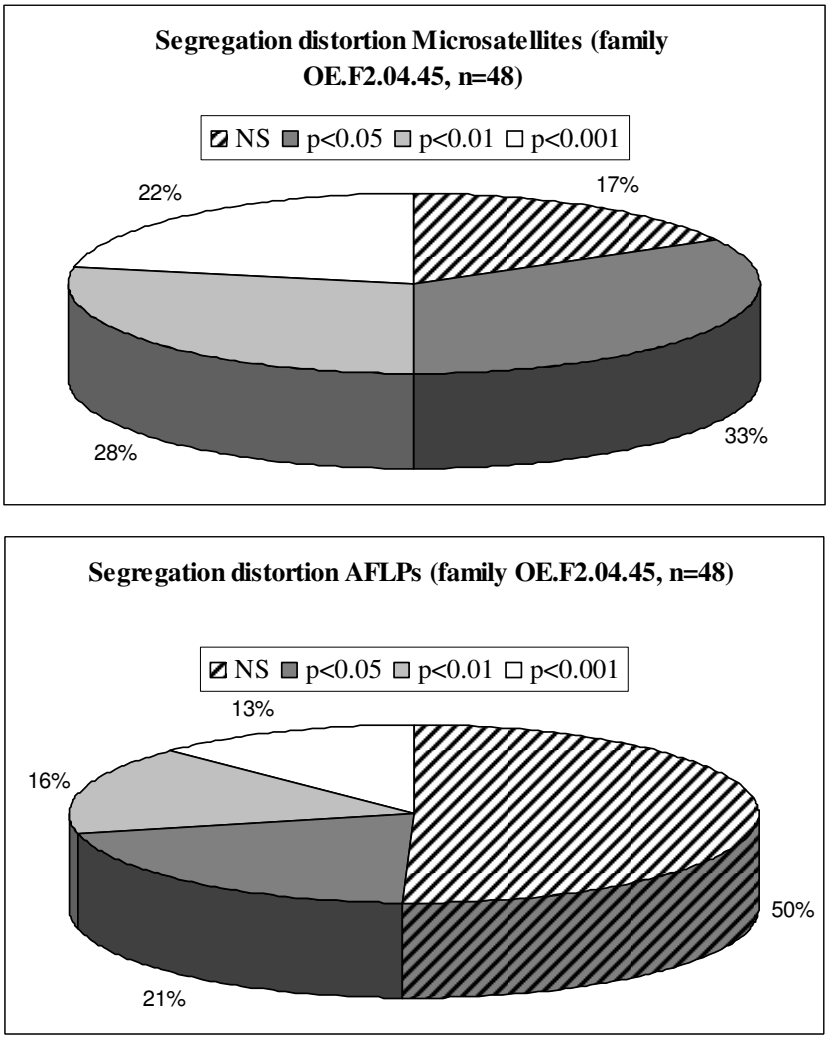


Figure 51. Segregation distortion for the microsatellite and AFLP markers for the mapping family OE.F2.04.45. NS: percentage of Mendelian markers; p<0.05, p<0.01, p<0.001: percentage of distorted markers at α=5%, 1% or 1‰.

For the second mapping family, OE.F2.04.63, 92 F2 progeny were genotyped for 20 microsatellites and 60 AFLP primer pairs. High segregation distortion was reported for this family, slightly less than for OE.F2.04.45. 25% of the microsatellites (4 out of 16 informative markers: Oe1/47, Oe3/37, *Oedu*.HA21 and *Oedu*.B11) were segregating according to Mendelian rules (Figure 52 a). Because this F2 family came from a real biparental cross, both 1:1 and 3:1 segregation types were reported for the AFLP markers. 61% of the 1:1 AFLPs (107 out of 175 markers) and 92% of the 3:1 AFLPs (65 out of 71 markers) were segregating according to Mendelian rules (Figure 52 b and c). Overall, 69% of the AFLP markers (172 out of 246 markers) were considered as Mendelian (Figure 52 d). The high percentage of Mendelian 3:1 AFLPs compared to the 1:1 AFLPs was due to the fact that a selection of 3:1 markers was done on the scoring of 48 F2 individuals only. In total, 251 markers were kept for further genetic linkage analysis, 16 microsatellites and 235 AFLPs (170 AFLPs type 1:1 and 65 Mendelian AFLPs type 3:1).

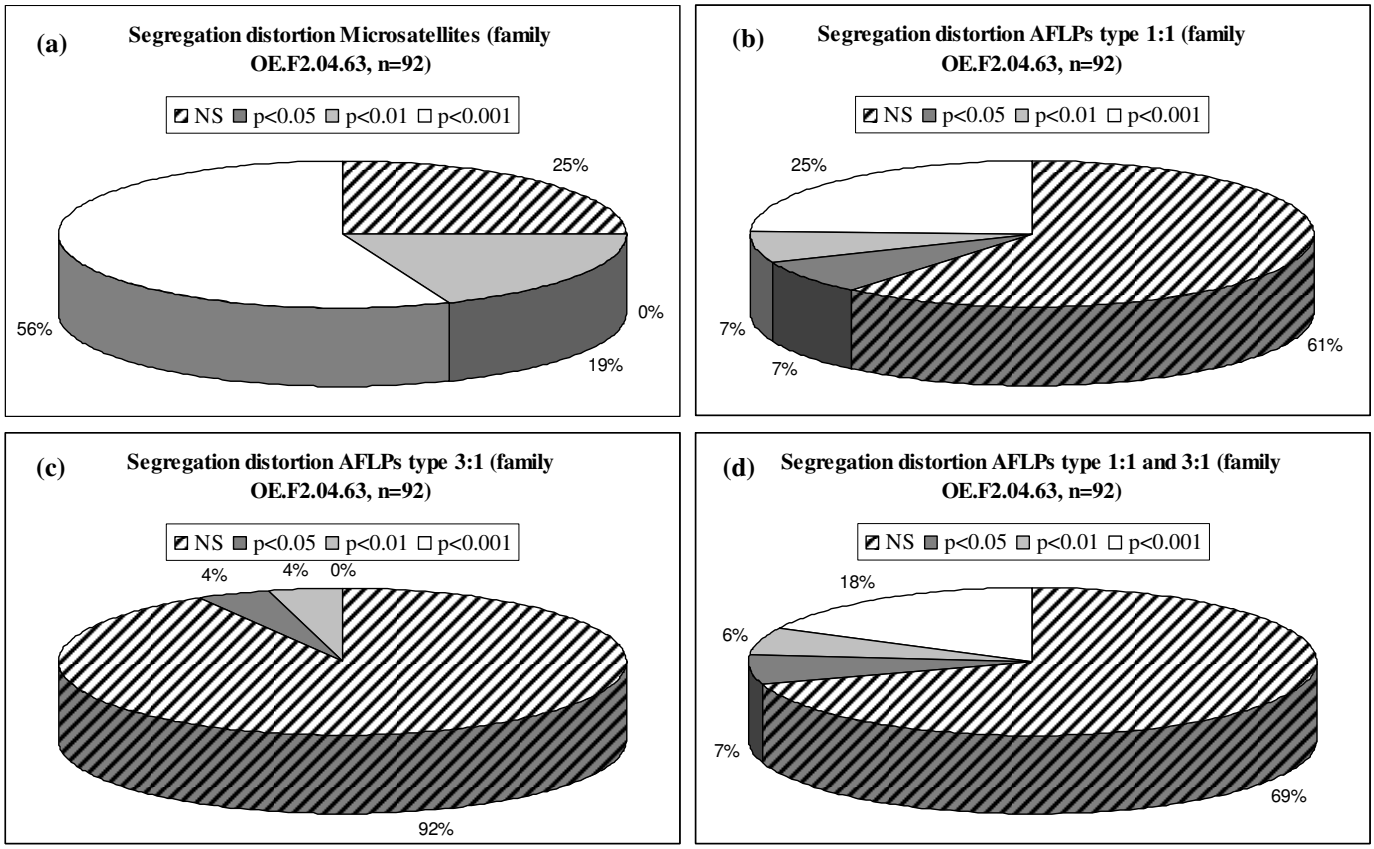


Figure 52. Segregation distortion for the microsatellite and AFLP markers for the mapping family OE.F2.04.63. NS: percentage of Mendelian markers; p<0.05, p<0.01, p<0.001: percentage of distorted markers at α=5%, 1% or 1‰.

IV.4. Genetic linkage analysis: family OE.F2.04.45

The first mapping family, OE.F2.04.45, was unusual and consisted of the two grand-parents (L002-55 and W120), one F1 parent (30-37) and 48 F2 progeny issued from a self-fertilisation of 30-37. Overall, 231 markers were kept for the genetic linkage analysis, 18 informative microsatellites and 213 Mendelian 3:1 type AFLPs.

Two-point analysis revealed 17 linkage groups, 10 with more than 8 markers, 3 triplets and 4 doublets. LOD scores were high, almost always higher than 4.0. However, most markers were linked with a recombination frequency of 0, making the ordering of markers within linkage groups very difficult. Three linkage groups containing 3 or 4 microsatellites lead to maps spanning 11.0, 26.8 and 49.9 cM. Microsatellites were ordered, and each microsatellite was associated to numerous AFLPs that were linked with a zero recombination frequency (Figure 53).

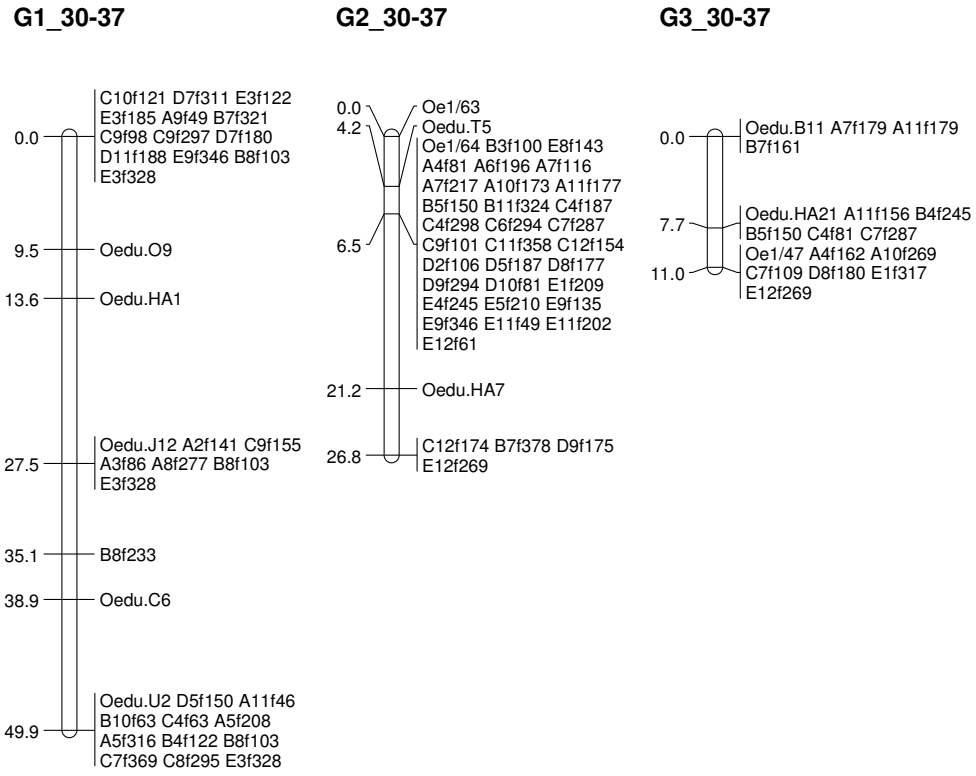


Figure 53. Three linkage groups containing 3 or 4 microsatellites for the mapping family OE.F2.04.45 obtained with CriMap software. AFLP markers are labelled with the primer pair name followed by the letter "f" for fragment and a 3-digit fragment size in base pairs. Markers are indicated on the right and absolute positions in Kosambi cM on the left. "Oedu.*" and "Oe*" correspond to microsatellites.

Four linkage groups containing one microsatellite spanned around 13-14 cM. They were characterised by one microsatellite distant from a huge cluster of AFLPs (containing 12 to 25 markers) (Figure 54).

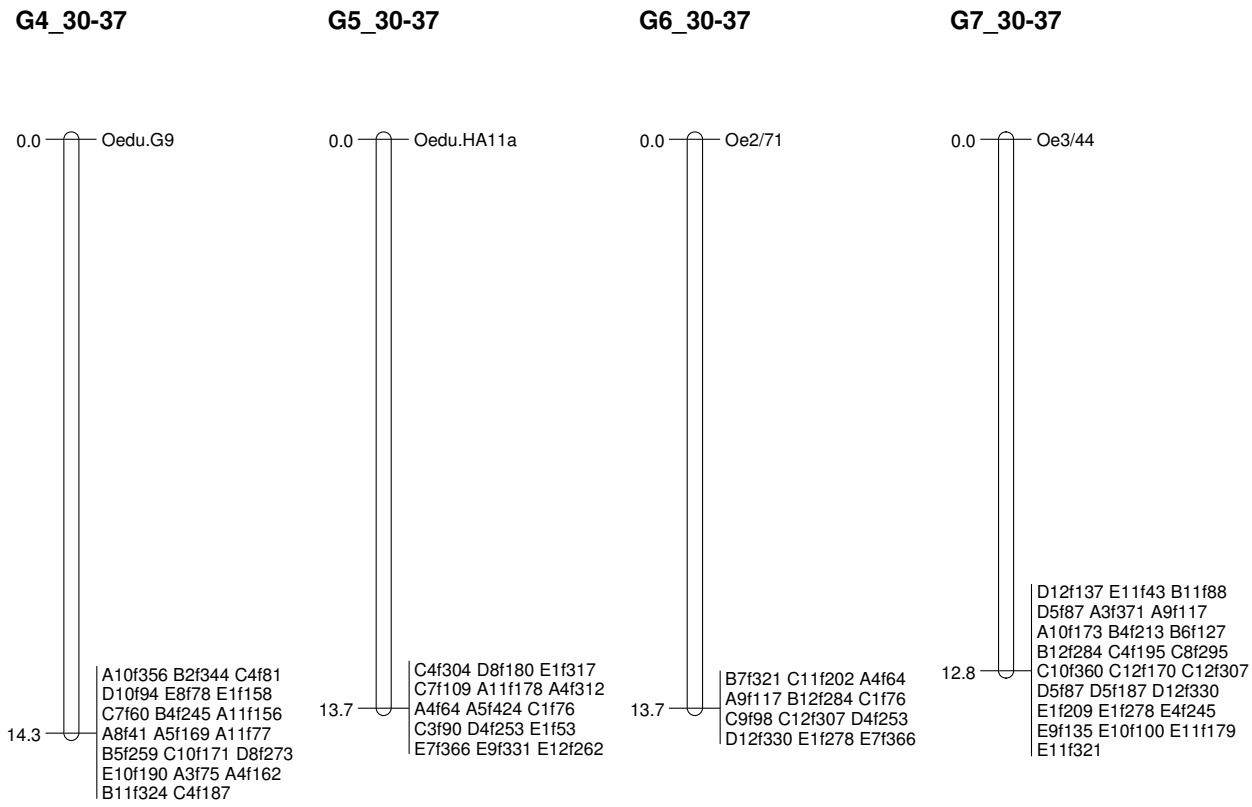


Figure 54. Four linkage groups containing 1 microsatellite for the mapping family OE.F2.04.45 obtained with CriMap software. AFLP markers are labelled with the primer pair name followed by the letter "f" for fragment and a 3-digit fragment size in base pairs. Markers are indicated on the right and absolute positions in Kosambi cM on the left.

The 10 other linkage groups did not contain microsatellite markers. Markers within linkage groups could not be ordered because of lack of informativeness of 3:1 AFLP markers and all AFLPs were linked with a zero recombination frequency. Three groups, G8_30-37, G9_30-37 and G10_30-37, contained from 8 to 11 AFLPs; three groups, G11_30-37, G12_30-37 and G13_30-37, contained three AFLPs; and the last four groups, G14_30-37, G15_30-37, G16_30-37 and G17_30-37, contained only 2 AFLPs (Figure 55).

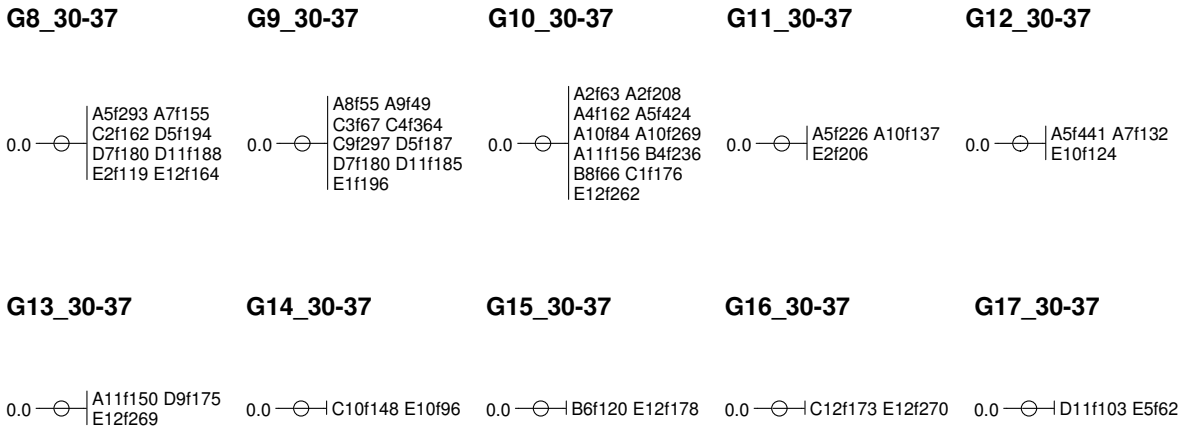


Figure 55. Ten linkage groups containing no microsatellite for the mapping family OE.F2.04.45, for which the ordering of markers within each linkage group was not possible due to the lack of informativeness of 3:1 AFLPs. CriMap software. AFLP markers are labelled with the primer pair name followed by the letter "f" for fragment and a 3-digit fragment size in base pairs. Markers are indicated on the right and absolute positions in Kosambi cM on the left.

IV.5. Sex-specific linkage maps: family OE.F2.04.63

The second mapping family, OE.F2.04.63, consisted of the two grand-parents (L002-53 and W102), the two F1 parents (23-31 and 23-32) and 92 F2 progeny. Two sex-specific maps were obtained with CriMap, one for the F1 parent 23-31 (later referred as "Parent 1" or "P1") and one for the F1 parent 23-32 (later referred as "Parent 2" or "P2").

The P1 genetic linkage map was based on the 16 microsatellites and 145 AFLPs segregating through this parent. The AFLPs consisted of 71 markers of type I (31 from L002-53, 40 from W102), 9 of type II and 65 of type III (37 from L002-53, 28 from W102) (see Figures 46 and 47 for reminder). The resulting map consisted of 104 markers (64.6%), comprising 14 microsatellites (from 16: 87.5%), 62 type I AFLPs (from 71: 87.3%), 7 type II AFLPs (from 9: 77.8%) and 21 type III AFLPs (from 65: 32.3%). Nine linkage groups were set up for the P1 map covering 471.2 cM (Figure 56). The sizes of the linkage groups ranged from 23.6 cM to 95.8 cM. The number of markers per linkage group varied from 4 to 22. The average distance between 2 loci ranged from 3.16 cM (G3_P1) to 10.1 cM (G8_P1), with an average spacing of 4.86 cM. The largest interval varied from 9.7 cM (G3_P1) to 35.3 cM (G4_P1) (Table 26).

Table 26. Length, number of markers, average spacing and largest interval of linkage groups of the Parent 1 (23-31) map in *O. edulis* established with CriMap. Family OE.F2.04.63.

Linkage group	Length (cM)	No. of markers	Average spacing (cM)	Largest interval (cM)
G1_P1	95.8	21	4.79	22.5
G2_P1	72.1	14	4.81	23.7
G3_P1	66.4	22	3.16	9.7
G4_P1	63.4	9	7.93	35.3
G5_P1	50.0	11	5.0	11.8
G6_P1	38.0	10	4.22	15.4
G7_P1	31.6	6	6.32	18.1
G8_P1	30.3	4	10.1	17.6
G9_P1	23.6	7	3.93	11.4
Total	471.2	104		

The P2 genetic linkage map was based on the 16 microsatellites and 154 AFLPs segregating through this parent. The AFLPs consisted of 84 markers of type I (38 from L002-53, 46 from W102), 5 of type II and 65 of type III (37 from L002-53, 28 from W102) (see Figures 46 and 47 for reminder). The resulting map consisted of 117 markers (76.0%), comprising 14 microsatellites (from 16: 87.5%), 76 type I AFLPs (from 84: 90.5%), 5 type II AFLPs (from 5: 100%) and 22 type III AFLPs (from 65: 33.8%). Ten linkage groups were set up for the P2 map covering 450.0 cM (Figure 57). The sizes of the linkage groups ranged from 11.9 cM to 77.7 cM. The number of markers per linkage group varied from 2 to 24. The average distance between 2 loci ranged from 1.34 cM (G8_P2) to 26.1 cM (G7_P2), with an average spacing of 4.21 cM. The largest interval varied from 10.5 cM (G4_P2) to 26.1 cM (G7_P2) (Table 27).

Table 27. Length, number of markers, average spacing and largest interval of linkage groups of the Parent 2 (23-32) map in *O. edulis* established with CriMap. Family OE.F2.04.63.

Linkage group	Length (cM)	No. of markers	Average spacing (cM)	Largest interval (cM)
G1_P2	77.7	15	5.55	23.0
G2_P2	69.8	10	7.76	24.2
G3_P2	68.8	24	2.99	17.8
G4_P2	65.4	17	4.09	10.5
G5_P2	50.7	13	4.23	18.9
G6_P2	45.8	11	4.58	14.2
G7_P2	26.1	2	26.1	26.1
G8_P2	17.4	14	1.34	11.0
G9_P2	16.4	6	3.28	12.1
G10_P2	11.9	5	2.98	10.8
Total	450.0	117		

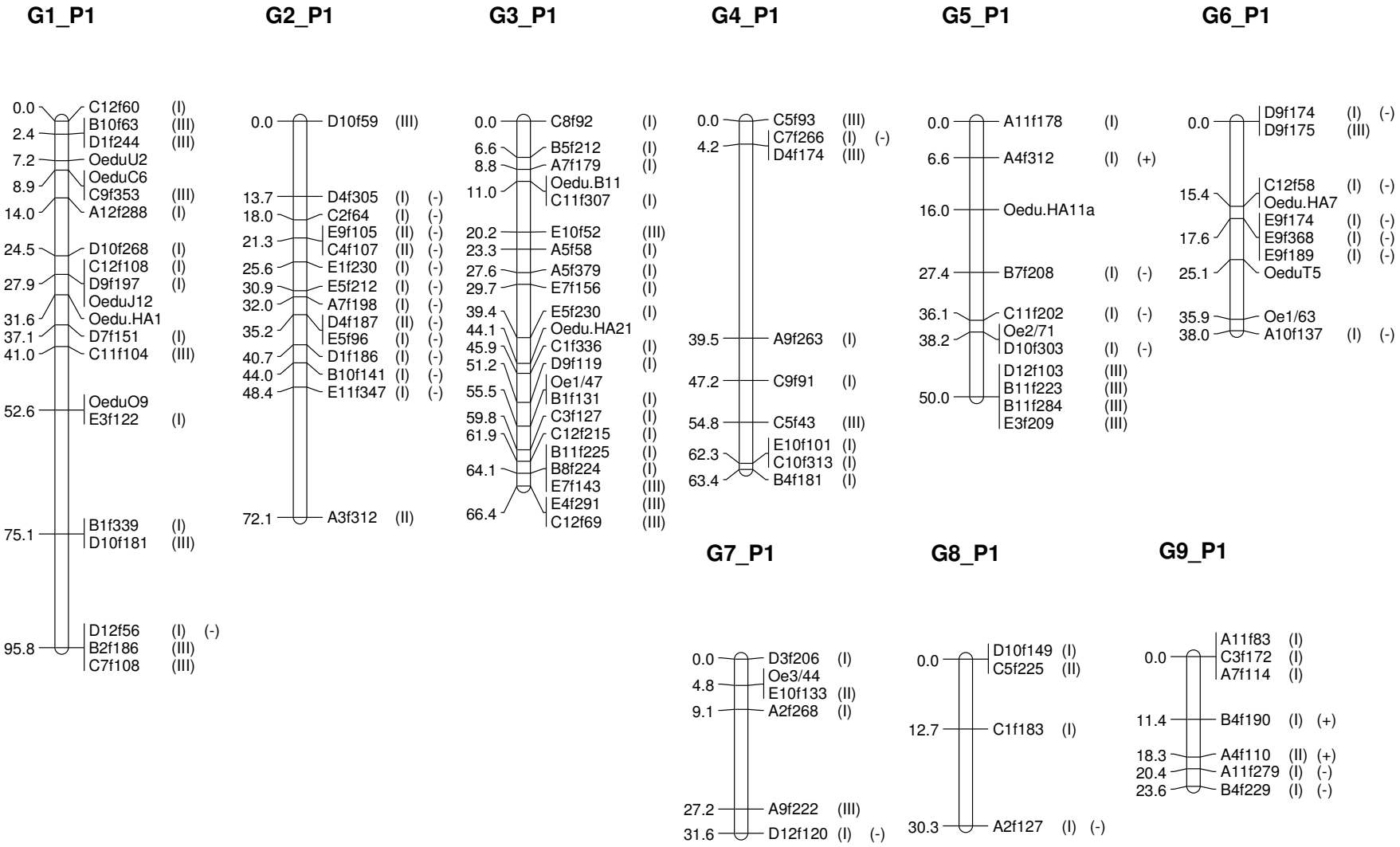


Figure 56. Microsatellite and AFLP-based linkage map of the flat oyster *O. edulis* in the mapping family OE.F2.04.63: P1 (23-31) map obtained with CriMap, 106 markers, 471 cM. AFLP markers are labelled with the primer pair name followed by the letter "f" (for fragment) and a 3-digit fragment size in base pairs. Markers are indicated on the right; and absolute positions on the left (in Kosambi cM). On the right of AFLP locus name are specified the segregation type (I, II or III; see Figures 46 and 47) and the direction of the segregation distortion: towards a deficit (-) or excess (+) of recessive homozygotes *aa*.

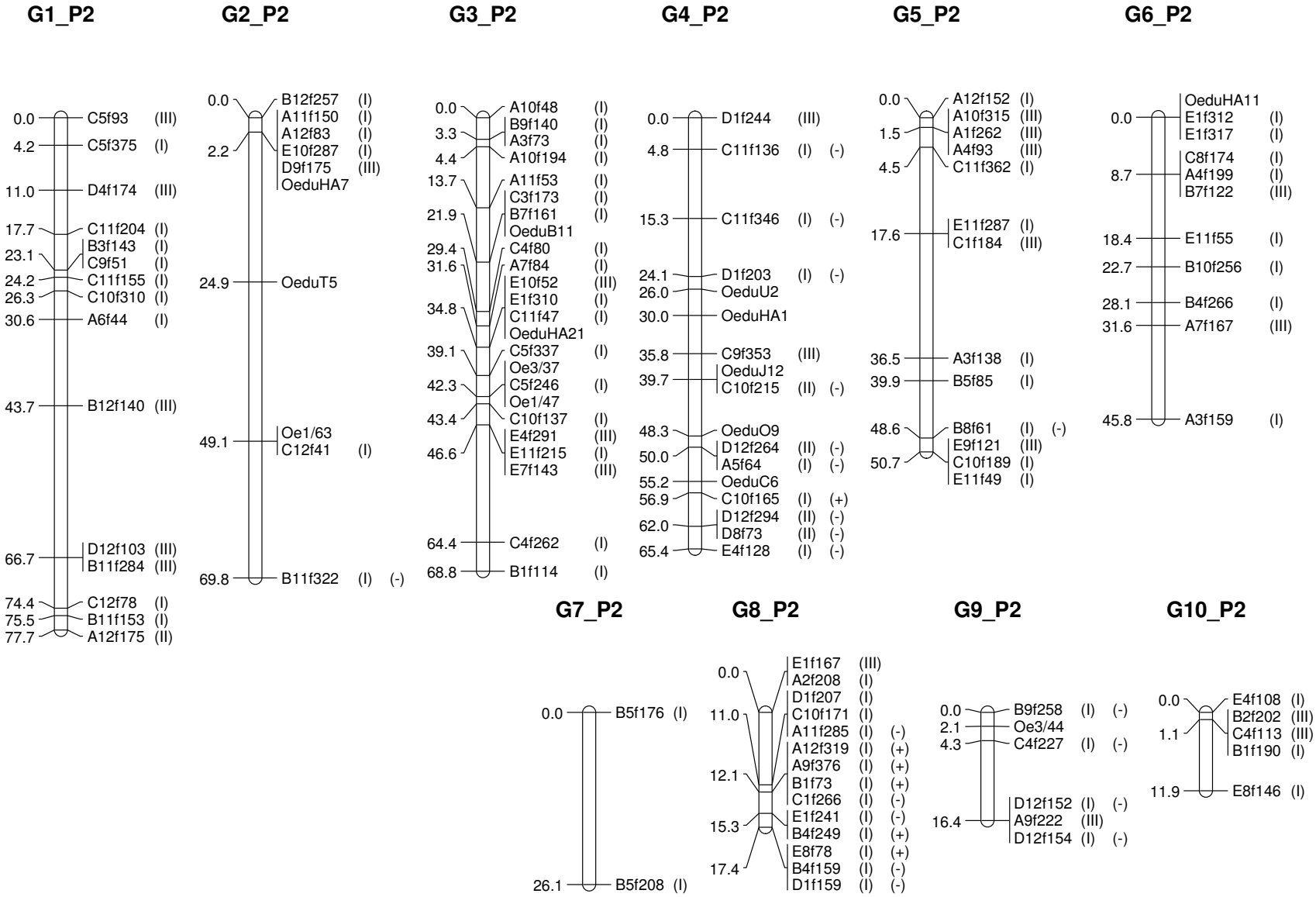


Figure 57. Microsatellite and AFLP-based linkage map of the flat oyster *O. edulis* in the mapping family OE.F2.04.63: P2 (23-32) map obtained with CriMap, 117 markers, 450 cM. See Figure 56 for abbreviations.

The observed map length was 471.2 cM for the P1 map. The estimated genome length was 571.7 cM according to method 4 of Chakravarti et al. (1991). The observed coverage was therefore 82.4 % for the P1 map. For the P2 map, the observed map was 450 cM and the estimated genome length 575.8 cM. The observed coverage was therefore 84.2 % for the P2 map.

Distorted markers showed non random distribution or clustering in both genetic maps (P1 and P2). In the P1 map, the 30 mapped distorted markers were mainly located on 4 linkage groups, G2_P1 (containing 12 homozygote deficiency markers in a 35 cM segment), G5_P1 (containing 3 homozygote deficiency markers and 1 homozygote excess marker), G6_P1 (containing 6 homozygote deficiency markers in a 38 cM segment) and G9_P1 (containing 2 markers with homozygote deficiency and 2 with homozygote excess). In addition, 4 other groups contained each only one distorted marker showing homozygote deficiency: G1_P1, G4_P1, G7_P1 and G8_P1 (Figure 56). In the P2 map, the 26 mapped distorted markers were mainly concentrated in 3 linkage groups, G4_P2 (containing 9 homozygote deficiency markers and 1 homozygote excess marker), G8_P2 (containing 5 markers with heterozygote deficiency and 5 with homozygote excess in a 5 cM segment) and G9_P2 (containing 4 homozygote deficiency markers in a 16 cM segment). In addition, 2 other groups, G2_P2 and G5_P2, contained each one distorted marker showing homozygote deficiency (Figure 57).

Distributions of loci from five EcoRI primers (Eco+CAG, Eco+ACG, Eco+AAC, Eco+ACA and Eco+ACT) and of microsatellites, over the linkage groups are shown in Figure 58 for both parental maps (P1 and P2). In the P1 map, primer Eco+CAG provided 18 loci mapping to 9 groups (ranging from 1 to 4 per group); primer Eco+ACG provided 14 loci mapping to 6 groups (ranging from 0 to 4 per group); primer Eco+AAC provided 24 loci mapping to 8 groups (ranging from 0 to 6 per group); primer Eco+ACA provided 20 loci mapping to 8 groups (ranging from 0 to 6 per group); primer Eco+ACT provided 16 loci mapping to 7 groups (ranging from 0 to 5 per group); and the 14 microsatellites mapped onto 5 groups (ranging from 0 to 5 per group) (Figure 58 a). In the P2 map, primer Eco+CAG provided 23 loci mapping to 8 groups (ranging from 0 to 5 per group); primer Eco+ACG provided 22 loci mapping to 9 groups (ranging from 0 to 4 per group); primer Eco+AAC provided 28 loci mapping to 9 groups (ranging from 0 to 7 per group); primer Eco+ACA provided 12 loci mapping to 5 groups (ranging from 0 to 5 per group);

primer Eco+ACT provided 18 loci mapping to 7 groups (ranging from 0 to 5 per group); and the 14 microsatellites mapped onto 5 groups (ranging from 0 to 5 per group) (Figure 58 b). However, a contingency chi-square test on this dataset could not be safely performed because there was not enough data in each expected class (less than 5) for accurate computation.

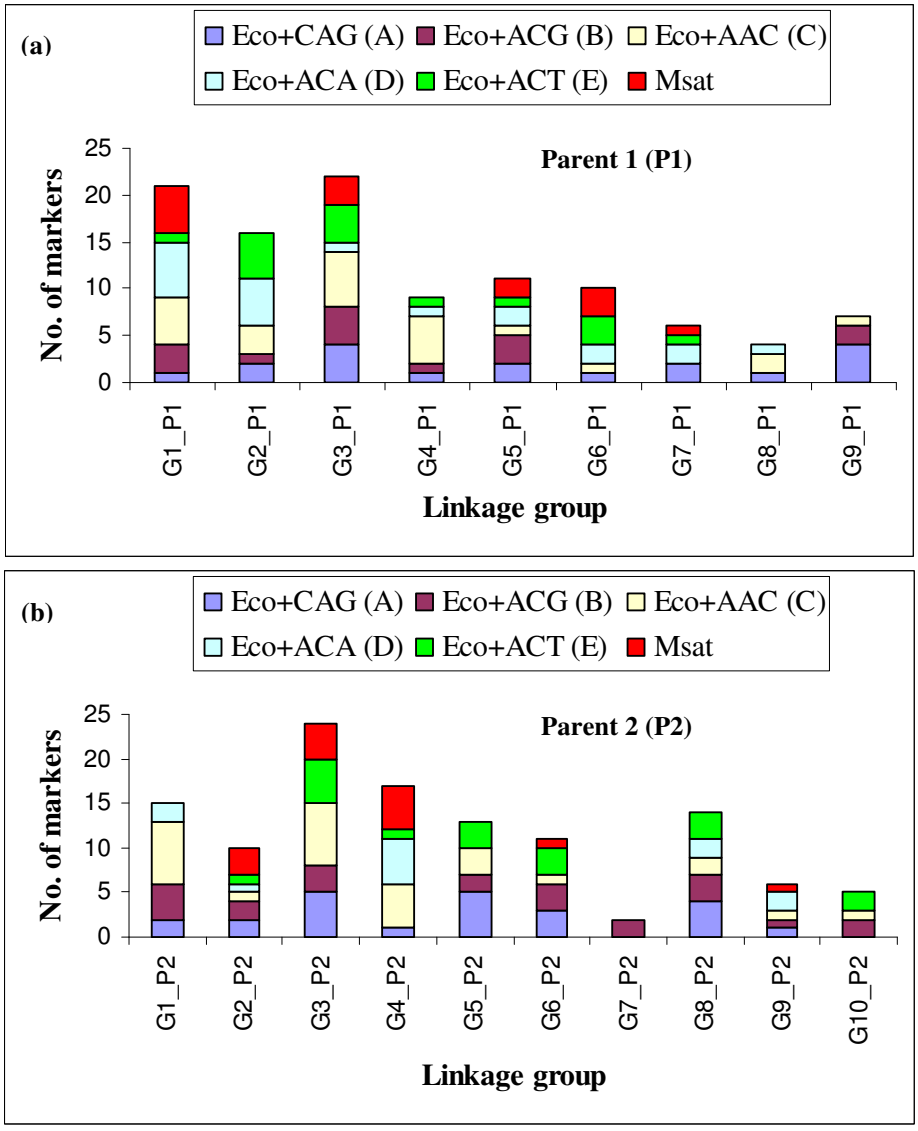


Figure 58. Histograms depicting the number of loci detected with 5 different labelled EcoRI primers and the number of microsatellites, for each linkage group of the P1 and P2 genetic linkage maps (family OE.F2.04.63).

Distribution of interval sizes between adjacent markers on the P1 and P2 maps revealed a high number of clusters, 34 for the parent 1 (23-31) and 54 for the parent 2 (23-32). In addition, a few gaps remained to be filled because there were 5 intervals spanning more than 18 cM in P1 and 6 in P2 (Figure 59).

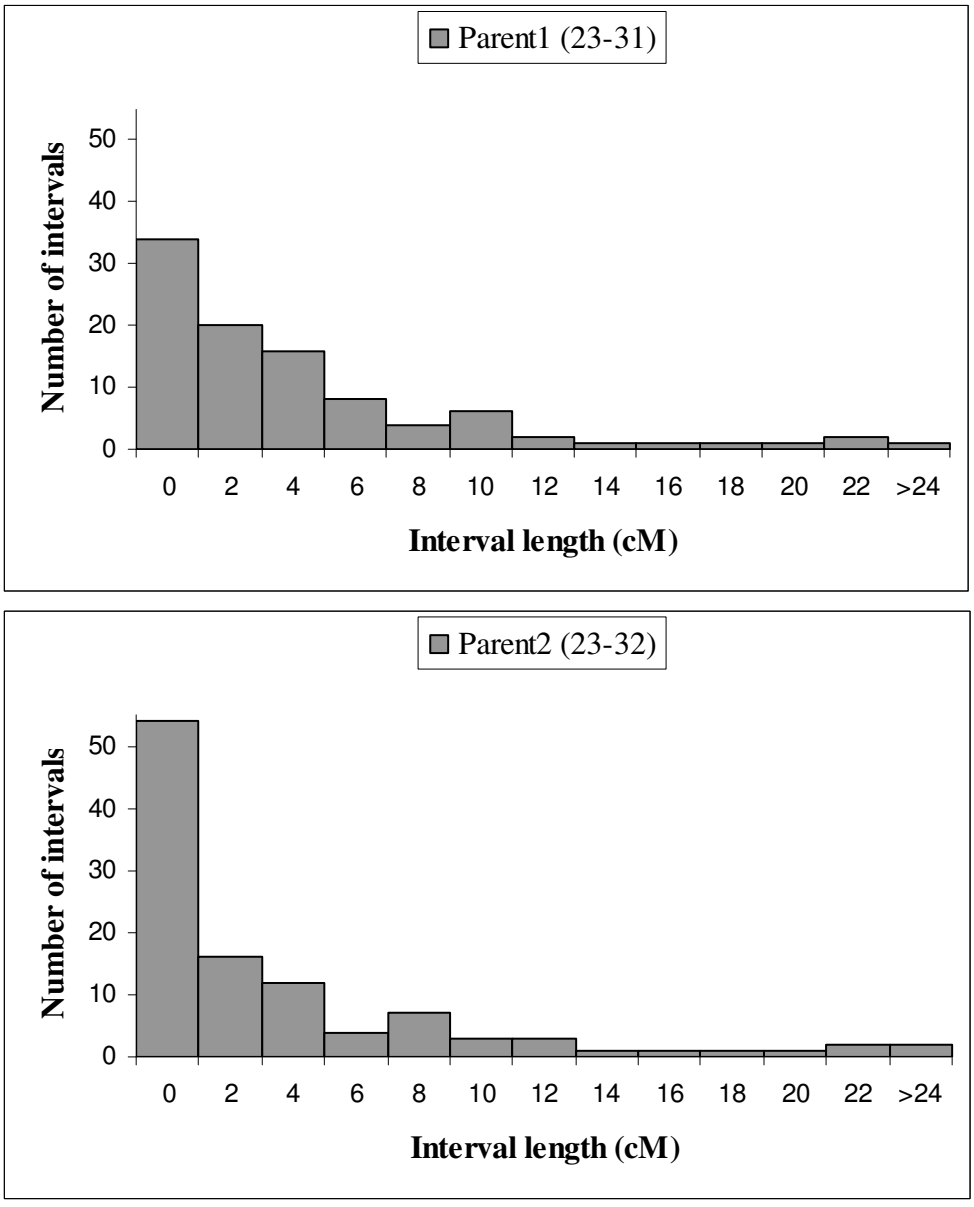


Figure 59. Distribution of interval sizes (in cM) between adjacent markers on the P1 and P2 maps of *O. edulis* (family OE.F2.04.63).

The assumption of a random distribution of AFLP markers in the genome was tested by Spearman correlation coefficients and chi-square test for departure from a Poisson distribution. Spearman correlation coefficients (r_s) between genetic length and number of markers per group were 0.85 for the parent 1 ($p<0.05$) and 0.61 ($p>0.05$) for the parent 2. Therefore, AFLP markers were randomly distributed in the linkage groups of the P1 map and not randomly distributed with the presence of some clustering of markers in the P2 map (Figure 60).

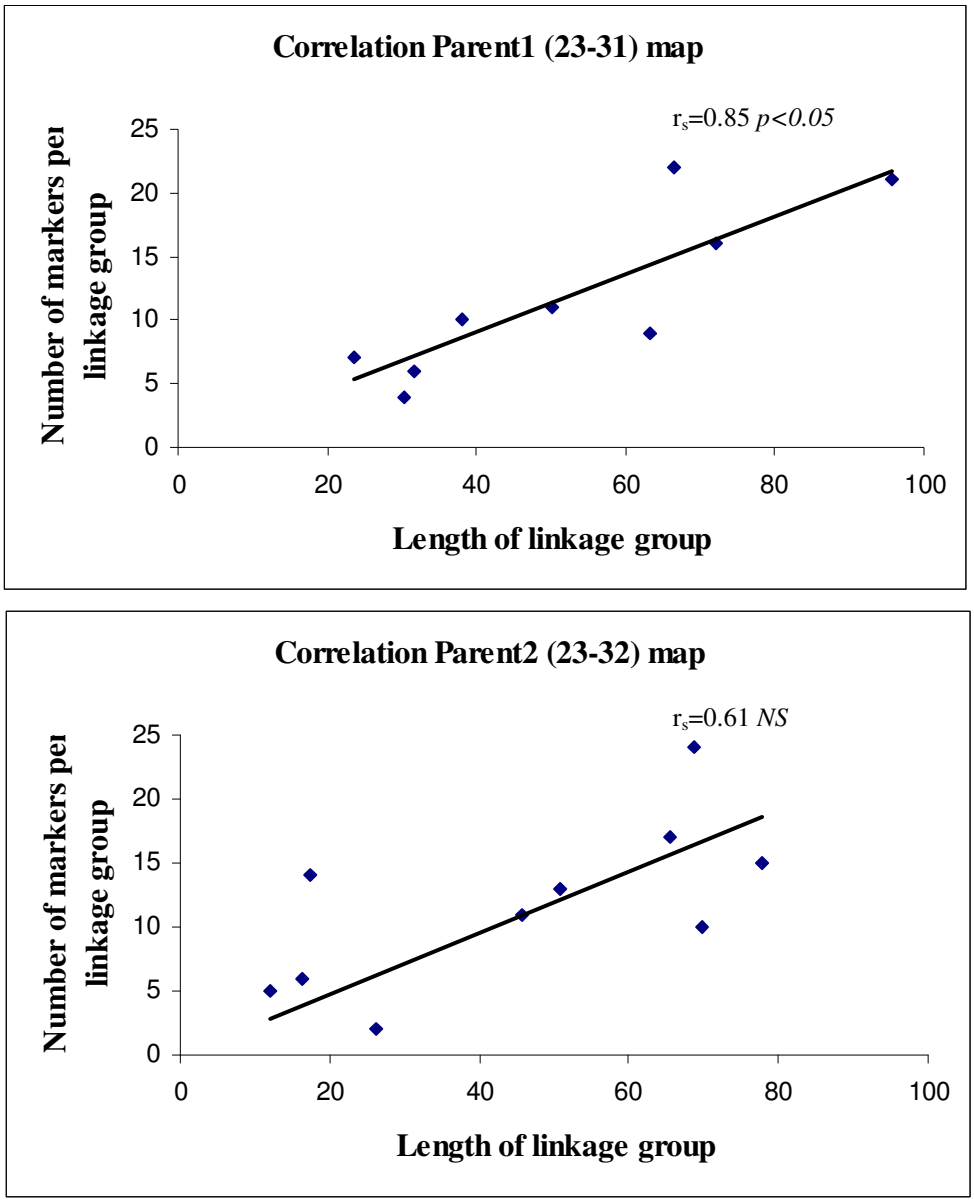


Figure 60. Correlation between length of linkage group (in cM) and the number of markers per linkage group in the P1 and P2 maps of *O. edulis* (family OE.F2.04.63). *p<0.05* corresponds to a significant correlation, implying that AFLP markers were randomly distributed in the genome. *NS* corresponds to an absence of correlation, implying a non-random distribution.

Observed and expected distributions of AFLPs and microsatellites were compared for 20 cM intervals in both P1 and P2 maps. A chi-square test for departure from a Poisson distribution was computed. The mean of the Poisson distribution was set up to 4.2 for P1 and 4.95 for P2. No significant departure from the Poisson distribution was observed for the parent 1 ($p=0.796$, 6 d.f.). However, this goodness-of-fit test was highly significant for the parent 2 ($p<0.001$, 6 d.f.), mostly due to 3 intervals of 20 cM containing only one microsatellite (large interval gaps remained to be filled in G1_P2, G2_P2 and G7_P2) and to one interval of 20 cM containing 14 markers (high clustering in G8_P2).

IV.6. Finding homology groups in the mapping family OE.F2.04.63

The mapping of microsatellite markers (codominant) and of type III AFLPs (segregating through both parent 1 and parent 2) in the two parental genetic linkage maps P1 and P2 allowed the finding of 8 probable homology groups. Homology groups were assumed when several markers (microsatellites or/and AFLPs) were linked in two linkage groups (P1 and P2) and when some of these markers were mapped in both P1 and P2 maps. For example, G1_P1 and G4_P2 exhibited 17 markers in common, including 5 microsatellites, and 7 of these 17 markers (including the 5 microsatellites) were mapped in both maps. Therefore, linkage groups G1_P1 and G4_P2 were assumed to be homology groups. Other homology groups were not supported by such confidence, such as the homology between G5_P1 and G6_P2 which relied on a single microsatellite (Table 28).

Table 28. Probable homology groups between P1 and P2 maps of the mapping family OE.F2.04.63, in *O. edulis*. On the same line are shown the names and lengths of the 2 homology groups between maps P1 and P2, the number of markers in common (belonging to those 2 linkage groups), and the number of markers mapped in both maps. In brackets are given the number of microsatellites. In two cases, markers in common were mapped only in the P2 map.

P1 group name	Length group P1 in cM	P2 group name	Length group P2 in cM	No. markers in common (msats)	No. markers mapped (msats)
G1_P1	95.8	G4_P2	65.4	17 (5)	7 (5)
G2_P1	72.1	G5_P2	50.7	4 (0)	4 in P2 only
G3_P1	66.4	G3_P2	68.8	7 (3)	7 (3)
G4_P1	63.4	G1_P2	77.7	3 (0)	2 (0)
G5_P1	50.0	G6_P2	45.8	1 (1)	1 (1)
G6_P1	38.0	G2_P2	69.8	4 (3)	4 (3)
G7_P1	31.6	G9_P2	16.4	7 (1)	2 (1)
G8_P1	30.3	G10_P2	11.9	2 (0)	2 in P2 only

After finding homology groups it was necessary to compare the order of markers mapped in both P1 and P2 maps in order to assess the accuracy of the maps obtained. This was possible for the 6 homology groups in which common markers were mapped in both P1 and P2 maps (Figure 61).

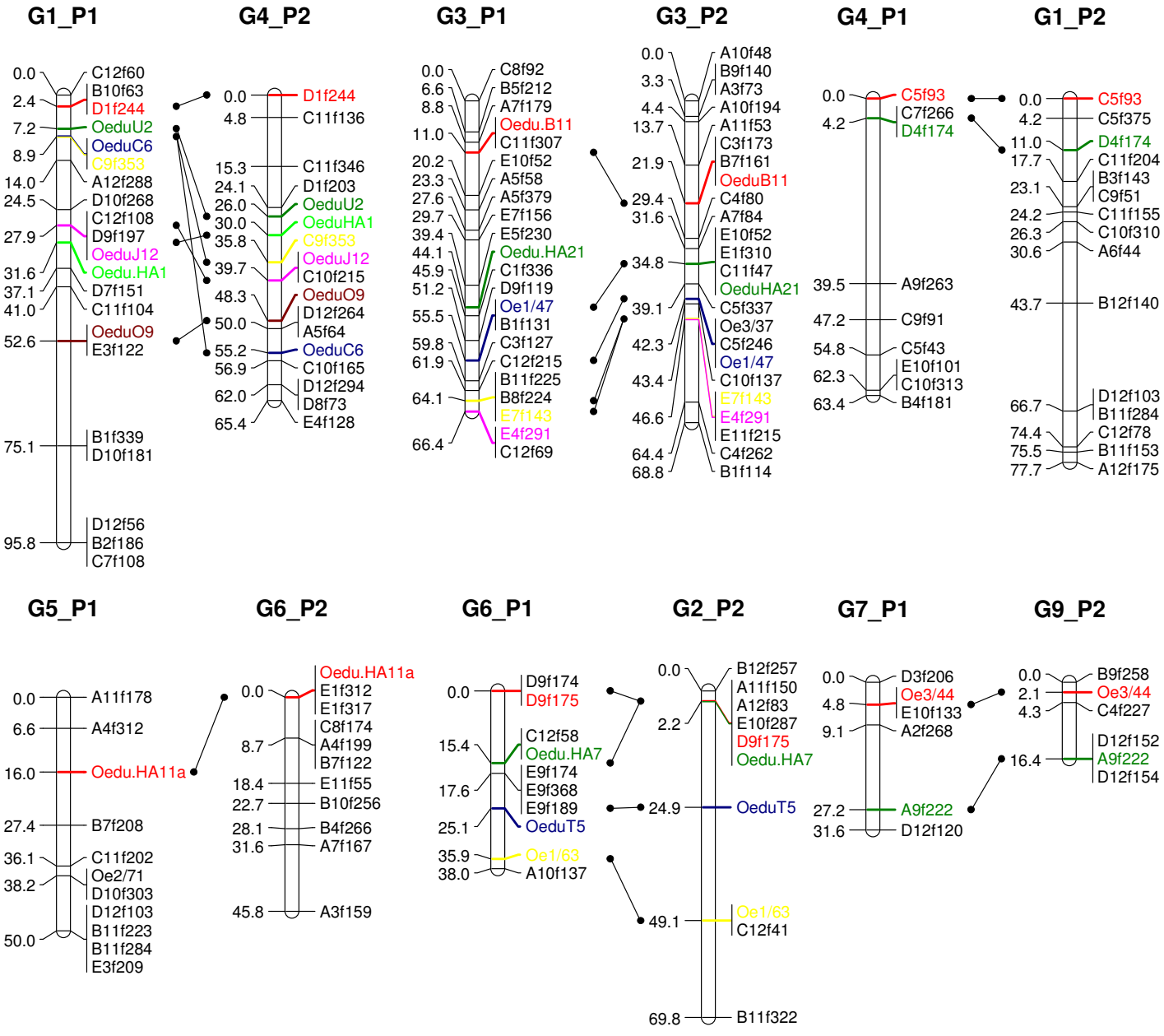


Figure 61. Six homology groups between P1 and P2 maps of the mapping family OE.F2.04.63 and comparison of markers orders. Six of the 8 homology groups are shown, for which common markers were mapped in both parental maps P1 and P2, allowing the comparison of markers orders. Lines link common markers between two homology groups and common markers of two homology groups have the same colour. See Figure 56 for abbreviations.

For the two other homology groups (G2_P1/G5_P2 and G8_P1/G10_P2), common markers could be mapped only in the P2 map (Table 28); in P1 those markers led to an ambiguous map position and were then discarded: they were linked but could not be mapped.

Marker orders seemed conserved for most homology groups. The most discrepancy occurred for the homology pairs G1_P1 and G4_P2 for which the 5 microsatellites were not mapped in the same order in both maps. In some instances, lengths of homology groups were similar (G3_P1 and G3_P2, G5_P1 and G6_P2). However, in most cases, there was a discrepancy in map lengths between the two maps P1 and P2. This was striking for G1_P1 / G4_P2 and G6_P1 / G2_P2. Therefore, recombination frequencies obtained in the two parents were plotted in order to assess eventual sex-specific recombination fraction differences. Differences in recombination frequencies were observed between the two parents. Eighty six pairs of markers were segregating in both parents (17 pairs of microsatellites and 69 pairs of microsatellite/AFLP). Forty four pairs of markers showed statistically different recombination frequencies between Parent 1 and Parent 2, 12 of these were associated with higher recombination in Parent 2 and 32 with higher recombination in Parent 1 (Figure 62).

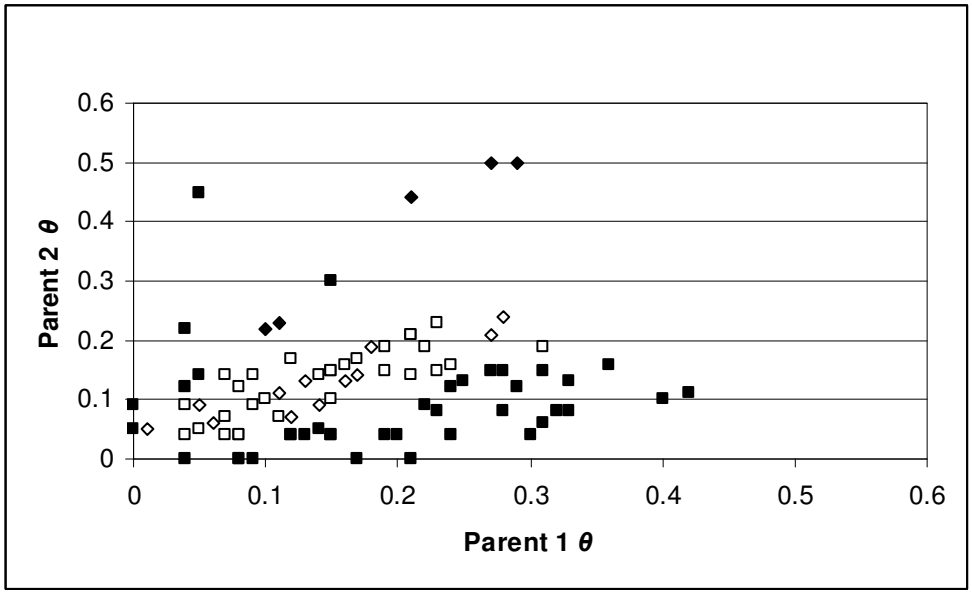


Figure 62. Parent 1 versus Parent 2 recombination fraction, for 86 pairs of markers segregating from both parents, in the mapping family OE.F2.04.63. Recombination fractions (θ) between 2 microsatellite markers (diamond) or between one AFLP marker and one microsatellite (square) were plotted. Open symbols are cases in which θ was statistically homogeneous between the two parents; solid symbols are cases in which θ was significantly heterogeneous between the two parents ($p<0.05$).

V- DISCUSSION

V.1. Self-fertilisation in *O. edulis*

Due to the particular reproductive features of *O. edulis* (brooding of eggs and early larvae in the mantle cavity of the females), the production of segregating families required holding pairs of oysters in small aquariums, with an outlet pipe for the collection of late larvae (Figure 40). It was assumed that the larvae collected resulted from the biparental cross between the two oysters held in the same tank because each spawning occurred at least one month after putting the oysters together (the brooding period in the mantle cavity of the female lasts around 10 days). However, parentage analysis (based on genotyping of microsatellites in the 2 parents and their offspring), performed a posteriori, revealed that the family OE.F2.04.45 came in fact from a self-fertilisation of one of the two oysters held in the same aquarium (Figure 50). This result was unexpected and surprising because to our knowledge no evidence of self-fertilisation has yet been reported in *O. edulis* in the literature. It is known that *O. edulis* is a sequential protandrous hermaphrodite species with possibility of changing of sex several times in the same reproductive season (Yonge, 1960; Le Dantec and Marteil, 1976). Moreover, parentage analyses performed on several successive mass spawnings from a pool of progenitors held in the same tank revealed the ability of flat oysters to change sex very quickly, in less than 2 weeks (Lallias et al., in prep). Such a quick sex change supposed that maturation of the two types of gametes (ovocytes and spermatozoids) was simultaneous, with probably a small delay in time in the maturity of the two types of gametes. Therefore, the co-occurrence of female and male gametes at the same time in the same individual confirms the possibility of self-fertilisation in *O. edulis*, at least in experimental conditions.

V.2. Segregation distortion: implications of the high genetic load

High segregation distortion was evident in the two segregating families of *O. edulis* analysed. For the family OE.F2.04.45 (n=48), 51.1% of the markers were significantly distorted, 83% of the microsatellites (15 out of 18) and 50% of the AFLPs (213 out of 424) (Figure 51). For the family OE.F2.04.63 (n=92), 32.8% of overall markers were distorted with microsatellites

being commonly so (75%) and AFLPs much less so (31%) (Figure 52). The range of segregation distortion reported in this study for the family OE.F2.04.63 issued from a real biparental cross was similar or slightly higher than that reported in another oyster species, *C. gigas*: 31% with allozymes (McGoldrick and Hedgecock, 1997); 20.9% with microsatellites (Launey and Hedgecock, 2001); 26.9% with AFLPs (Li and Guo, 2004). However, the family derived from the self-fertilisation of an F1 parent (OE.F2.04.45) exhibited higher segregation distortion.

In our study, the high proportion of distorted AFLP markers that showed an *aa* homozygote deficiency could be explained by a high genetic load that has previously been reported in bivalves (McGoldrick and Hedgecock, 1997; Bierne et al., 1998; Launey and Hedgecock, 2001). The two mapping families (OE.F2.04.45 and OE.F2.04.63) originated from a 6^{th} generation inbred line (6 generations of full-sib matings) that would certainly have undergone some inbreeding depression. Therefore, assuming that purging of deleterious genes by full-sib crosses was not complete by the 6^{th} generation, it is likely that the high segregation distortion observed was due to linkage of markers with lethal or deleterious genes in the recessive state.

Moreover, the distribution of distorted markers was not random in the genetic linkage maps produced. Clusters of distorted markers were restricted to a few linkage groups (Figures 56 and 57). Several studies reported that the mapping of distorted markers may help to understand the distribution of deleterious recessive genes in the genome. Indeed, Yu and Guo (2003) reported in *C. virginica* the clustering of 6 distorted markers in an area spanning 5 cM, which could potentially correspond to the presence of a deleterious gene nearby. In the same way, Li and Guo (2004) reported the mapping of at least 4 major deleterious recessive genes in the female map of the Pacific oyster *C. gigas*, highlighted by the clustering of markers with segregation distortion in the same direction. In the rainbow trout (*Oncorhynchus mykiss*) map, large regions of linkage groups contained blocks of distorted markers that could be linked to sub-lethal genes (Young et al., 1998). In our study, distorted markers tended to cluster on specific linkage groups and sometimes to a small segment of a linkage group (spanning 5 to 15 cM). These clusters of distorted markers could therefore correspond to the location of potential deleterious genes in *O. edulis*.

V.3. Linkage map and genome coverage

This study presents the first genetic linkage maps for the European flat oyster *O. edulis* and the first linkage maps in any flat oyster species. Genetic linkage maps have been established for almost all major aquaculturally important species, including tilapia (e.g. Agresti et al., 2000), catfish (e.g. Waldbieser et al., 2001), salmon (e.g. Moen et al., 2004b), rainbow trout (e.g. Nichols et al., 2003), abalone (e.g. Baranski et al., 2006) and shrimps (e.g. Li et al., 2003). However, it is striking that the studies concerning bivalve species including oysters are very scarce. A few studies reported the construction of a preliminary genetic map in *C. gigas* (Hubert and Hedgecock, 2004; Li and Guo, 2004), in *C. virginica* (Yu and Guo, 2003) or in *C. farreri* (Li et al., 2005; Wang et al., 2005).

The genome coverage achieved in *O. edulis* was good, above 82% for the P1 and P2 parental maps (family OE.F2.04.63). Those genome coverage estimates compared favourably with the ones established in cupped oysters' species which were in the range of 70-90% depending on the study (Yu and Guo, 2003; Hubert and Hedgecock, 2004; Li and Guo, 2004). Moreover, the number of linkage groups in P2 matched the haploid number of 10 chromosomes in this species (Thiriot-Quiévreux and Ayraud, 1982; Thiriot-Quiévreux, 1984b) although only 9 linkage groups could be clearly identified in P1 (Figures 56 and 57). This discrepancy for P1 suggests that gaps remained to be filled and that more markers should be added to the maps for a better coverage of the genome. This was confirmed by the fact that only 8 probable homology groups could be found in the mapping family OE.F2.04.63 (Figure 61). Finally, some of the linkage groups consisted of only two markers or spanned a small genetic distance (< 20 cM). Therefore, these groups may in fact belong to the same chromosome and may coalesce by adding more markers.

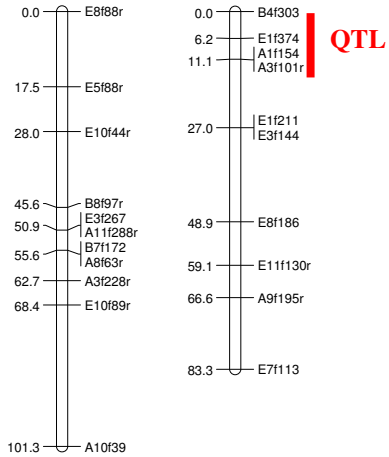
V.4. Recombination differences between the sexes

In the mapping family OE.F2.04.63, higher recombination rates in the parent P1 were reported, with 32 significant pairwise recombination rate differences (out of 44) compared with parent P2 (12 out of 44 significant pairwise comparisons) (Figure 62). Unfortunately, because of

the brooding behaviour of *O. edulis* the sex of our individual F1 oyster parents could not be determined. It could be of interest to sequence some mitochondrial fragments of the two parents and a few offspring to determine which parent was female (maternal transmission). Large sex-specific differences in recombination rates have been reported in several studies. Higher recombination rates in females were found in rainbow trout (Sakamoto et al., 2000), channel catfish (Waldbieser et al., 2001), zebrafish (Knapik et al., 1998), *C. virginica* (Yu and Guo, 2003) and the black tiger shrimp *Penaeus monodon* (Wilson et al., 2002). Therefore, these potential sex-specific differences in recombination rates in *O. edulis* should be confirmed and investigated further by mapping more anchor markers.

CHAPTER 5

Identification of QTLs for *Bonamia ostreae* resistance in *Ostrea edulis*



I- INTRODUCTION

Since 1985, Ifremer (Institut francais de recherche pour l'exploitation de la mer), which is a French research institute for exploration of the sea, has been undertaking a selective breeding programme for resistance to the protozoan parasite *Bonamia ostreae* (Haplosporidian protist, Carnegie et al., 2000) with the main aim of producing families of flat oysters tolerant to bonamiasis which could be compatible with the re-establishment of the cultivation of this species. Two improved oyster strains (S85 and S89) were produced by mass spawning (Naciri-Graven et al., 1998). However, such a character is difficult to select for as (1) the pressure of selection cannot be controlled (neither in the field nor in the laboratory where the parasite cannot be reproduced), (2) the resistance to this parasite is usually expressing at the third summer of life of the oysters, leading to a long generation time in this selective breeding programme.

Moreover, it has been shown in Chapter 1 that three-generation pedigree (grand-parents, parents and progeny) is a commonly used and powerful experimental design for QTL mapping. Such segregating families were produced. Clearly the establishment of a genetic linkage map for the flat oyster is an important step towards the understanding of genetic component of resistance/susceptibility to *B. ostreae* by the identification of QTLs of resistance to bonamiasis and the eventual development of marker-assisted selection. Moreover, the two improved oyster strains (S85 and S89) represent a valuable resource for genetic mapping in the context of QTL mapping. MAS would represent a critical tool in the acceleration of the selective breeding process by reducing the time between two generations of selection (3 years otherwise).

The aim of this part of the work was to find QTLs of resistance or susceptibility to bonamiasis in the European flat oyster *Ostrea edulis*. This relied on a 6-month trial challenge experiment in which wild oysters (overinfected with *B. ostreae* after purification of the parasite) were put in cohabitation with the tested oysters (segregating family), meaning that wild oysters and the tested oysters were held in the same tank. Mortality was checked daily. The aim of this experiment was to allow transmission of bonamiasis from wild oysters (overinfected with the parasite) to the tested oysters, and to link genotypes (different alleles of molecular markers) to phenotypes (rapidity of death, level of infection to the parasite).

II- PROGRAMME OF SELECTION FOR RESISTANCE TO BONAMIASIS (IFREMER)

II.1. Aquaculture production of the European flat oyster *O. edulis*

Within the past forty years, global production of *O. edulis* showed a drastic decline from a peak output of nearly 30,000 tons in 1961, due to the impact of diseases and a consequential shift to the rearing of the Pacific cupped oyster (*Crassostrea gigas*). European flat oyster production has remained low throughout the decade 1993-2002; output peaked in 1996 (7,996 tons) but became more stable (6,000-7,000 tons) in 2000-2004 (Figure 63).

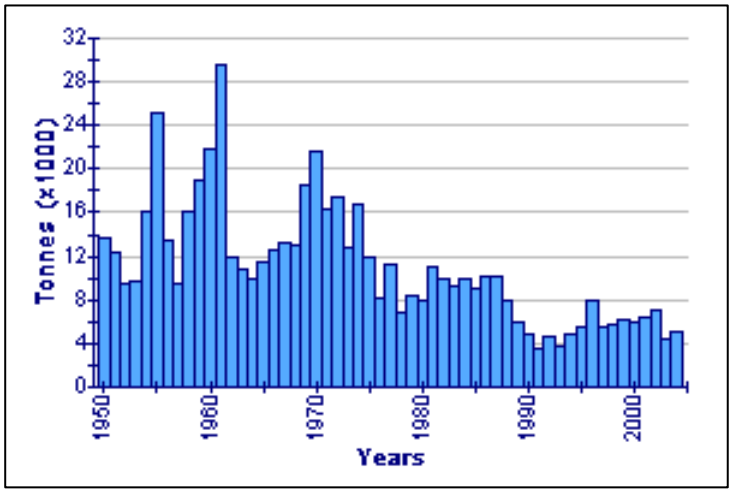


Figure 63. Global aquaculture production of the European flat oyster *O. edulis* (FAO Fishery Statistics).
http://www.fao.org/figis/servlet/static?dom=culturespecies&xml=Ostrea_edulis.xml

Almost all the production of *O. edulis* is in Europe. In 2002, 67 percent of the production was in Spain (4,565 tons) and 24 percent in France (1,600 tons). Ireland and the UK were the only other countries that produced more than 200 tons in 2002. The production of the European flat oyster constituted less than 0.2 percent of the total global production of all farmed oyster species in 2002. The bulk of production (97.7%) came from the rearing of the Pacific cupped oyster, *C. gigas*. However, the value of farmed *O. edulis* production in 2002 was 18.2 million euros (US$ 24.3 million); thus its culture remains an important sector in the limited areas where it is reared.

With regard to disease, massive mortalities widely struck European flat oyster populations in 1920. Some populations later recovered but many were replaced by two cupped oysters in several traditional rearing areas: the Portugese cupped oyster *Crassostrea angulata* in 1860 and the Pacific oyster *C. gigas* in 1970's after the drastic decline of the *C. angulata* production in the 1960's due to an iridovirus. Then, two parasitic diseases (due to *Marteilia refringens* and *B. ostreae*) spread in the late 1960's and 1970's, drastically reducing the production of *O. edulis* (Figure 63). This applied to most European countries, including France where the production decreased from 20,000 tons in 1950's and 1960's to around 2,000 tons nowadays (Figure 64).

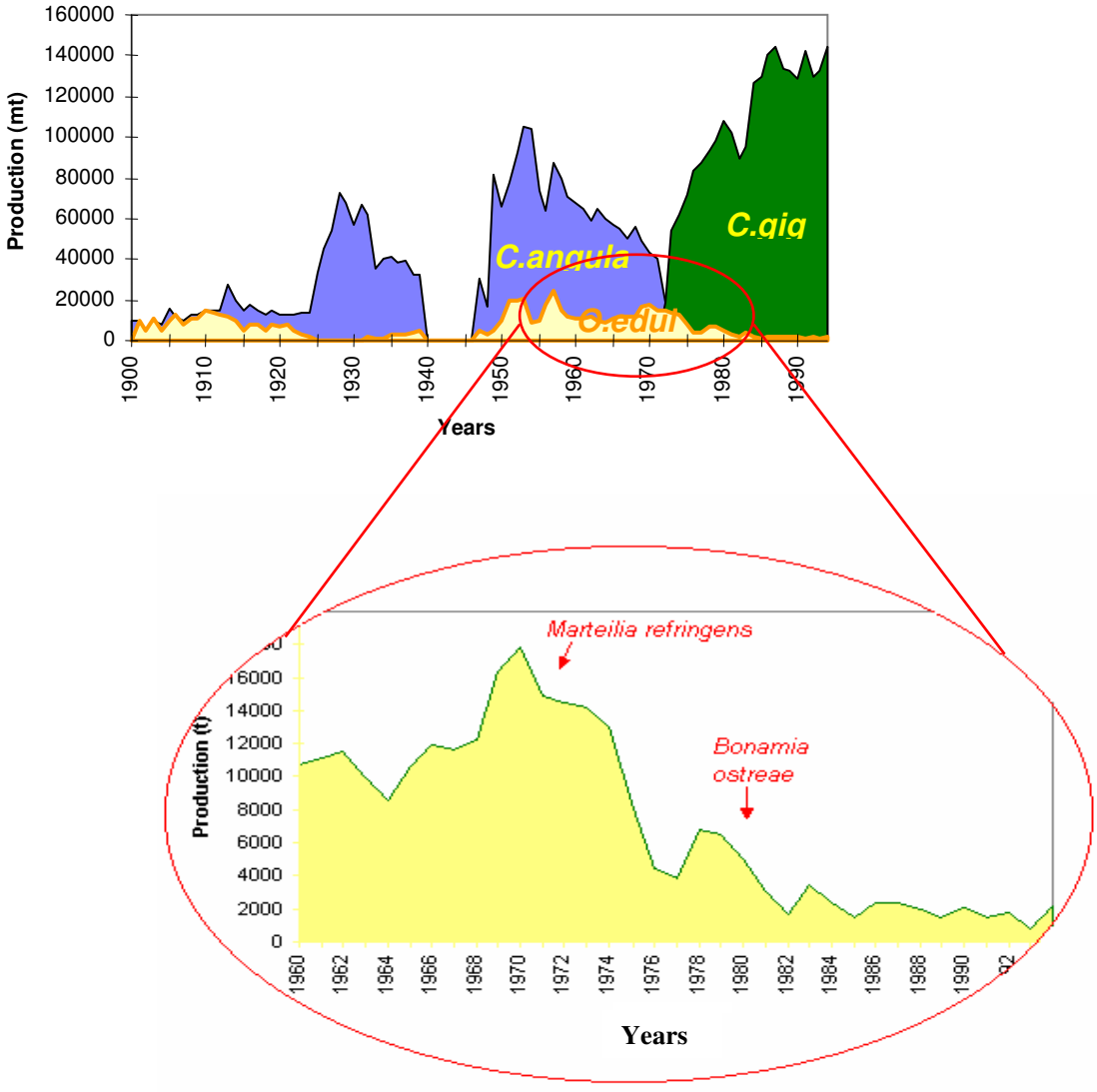


Figure 64. Oyster aquaculture production in France since the beginning of the 20th century (From Goulletquer and Heral, 1997).

II.2. Bonamiasis

II.2.1. Causative agent: Bonamia ostreae

Bonamiasis is an intrahaemocytic parasitosis due to the protist *B. ostreae* of the phylum Haplosporidia (Carnegie et al., 2000). Its morphology and ultrastructure has been studied by Comps et al. (1980) and Pichot et al. (1980). It is an intracellular parasite (2-4µm) affecting the haemocytes (blood cells) of the flat oyster (Figure 65). Two distinct cellular types can be observed:

- dense cells, highly basophilic, are the most frequently observed in the tissues
- clear cells, slightly basophilic, are more rare in the oyster tissues

Some plasmodial forms have been observed, especially in post-mortem oysters (Brehelin et al., 1982). This could explain the explosive development of the parasite before the death of an infected oyster, by allowing a quick multiplication of the parasite into the host (Launey, 1998).

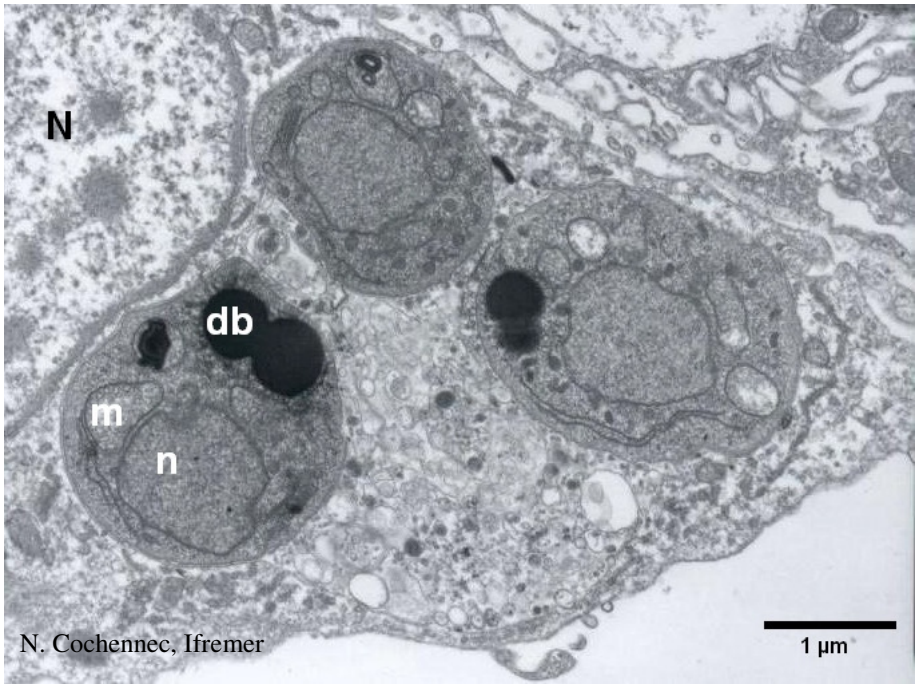


Figure 65. Three cells of *B. ostreae* in an haemocyte. N: nucleus of the host cell (haemocyte), n: nucleus of *B. ostreae*, m: mitochondria of the parasite, db: dense body.

II.2.2. Geographic distribution

B. ostreae was first discovered in Europe in Brittany (France) (Pichot et al., 1980; Comps et al., 1980). This disease, so far unknown from French production systems, spread very quickly. The parasite was then reported in other European countries: Netherlands (Van Banning, 1985, Van Banning, 1991), Spain (Figueras, 1991), Denmark (Elston et al., 1987), England and Ireland (Rogan et al., 1991) and in North America (Elston et al., 1986; Friedman et al., 1989; Friedman and Perkins, 1994). The geographical distribution of *B. ostreae* is illustrated Figure 66.

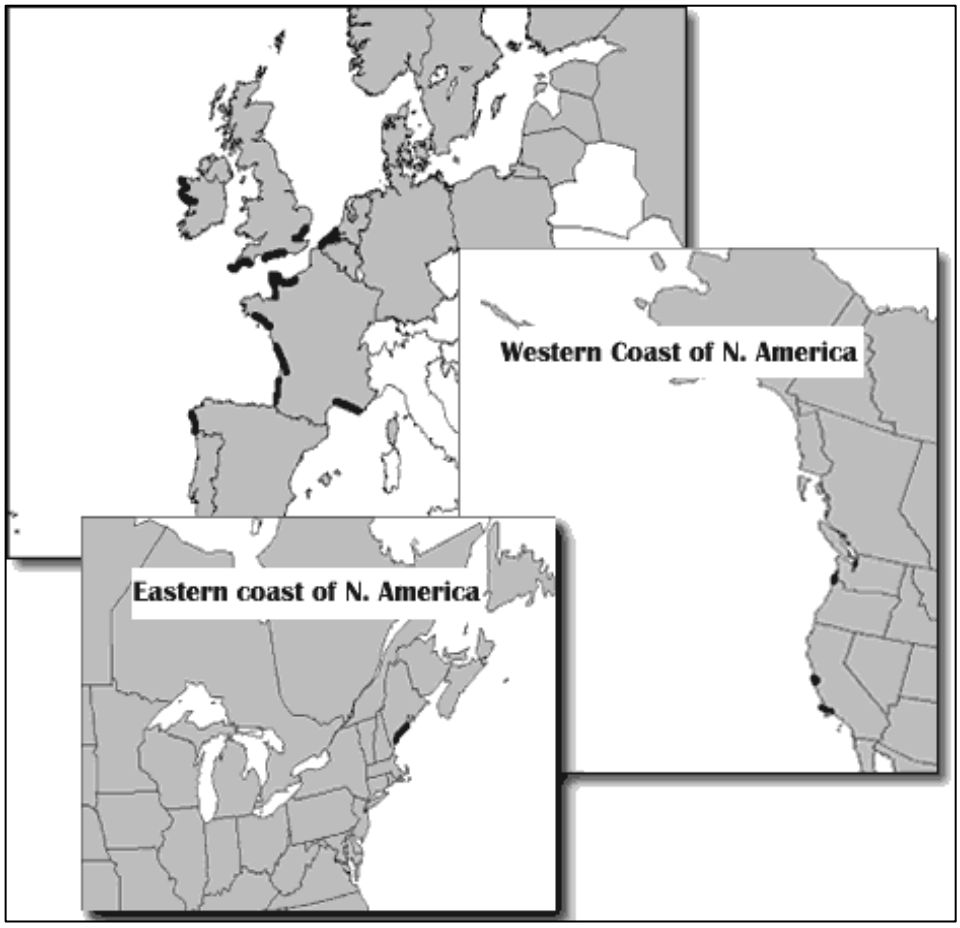


Figure 66. The geographical distribution of *B. ostreae* (heavy black line) in Europe and North America.
http://www.ices.dk/marineworld/fishdiseases/map8_5.htm

Current evidence suggests that *B. ostreae* originated in California and was inadvertently introduced into Maine, Washington and Europe by the translocation of infected *O. edulis* from California in the late 1970's (Elston et al., 1986, Friedman and Perkins, 1994, Cigarría and Elston 1997).

II.2.3. Pathology

Clinical signs of disease include a generally poor condition (no more growth) and gaping of oysters. Actual pathology appears correlated to haemocyte destruction and haemocytic infiltration of the connective tissue due to proliferation of *B. ostreae* (Balouet et al., 1983; Cochennec-Laureau et al., 2003). Lesions occur in the connective tissues of the gills, mantle, and digestive gland. Systemic infection of haemocytes effectively starves the oyster of energy required for survival.

In one study, the presence of *B. ostreae* was better related to size than to age of *O. edulis* and infection level was statistically independent of gonadal development stage (Caceres-Martinez et al., 1995). However, Robert et al. (1991) and Culloty and Mulcahy (1996) found that two years appeared to be the critical age for disease development in *O. edulis* in the Bay of Arcachon, France and on the south coast of Ireland, respectively. Nevertheless, young oysters (2-3 month-old to 18 month-old) were susceptible to infection and developed a high prevalence and intensity of infection over a six-month period (Lynch et al., 2005).

Significant mortalities usually occur at water temperatures of 12°-20°C. The disease can occur at any time of the year, although the prevalence and intensity of infection increase during the warmer months. Prevalence is generally highest in September. The pre-patent period is up to five months (Bower and McGladdery, 1996).

II.2.4. Diagnostic techniques

Three light microscopic techniques have been used so far for the detection of the parasite *B. ostreae*: tissue imprints, haemolymph cell monolayers and histological sections.

Tissue imprints correspond to smears from gill or heart tissue (preferably the ventricle since the auricles contain an abundance of serous cells which make detection of the parasite difficult), stained with Wright, Wright-Giemsa or equivalent stain (Figure 67).

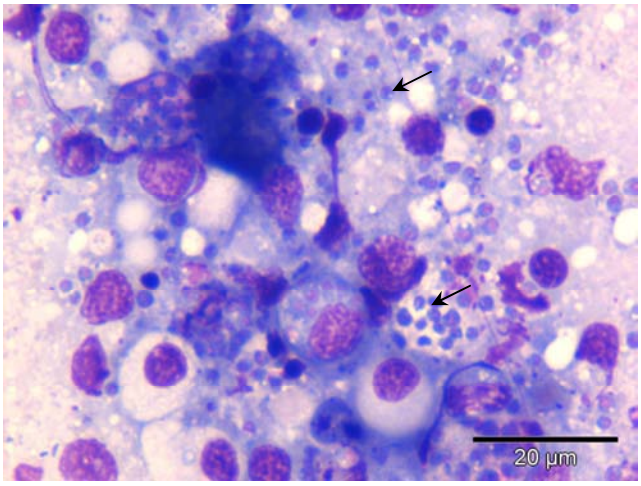


Figure 67. Hemacolor-stained heart smear achieved in *O. edulis*, allowing the detection of the parasite *B. ostreae* (visualized by arrows).

Histology consists of analysing haematoxylin and eosin stained tissue cross-sections and allows a better localisation of *B. ostreae* in the oyster tissues (Figure 68).

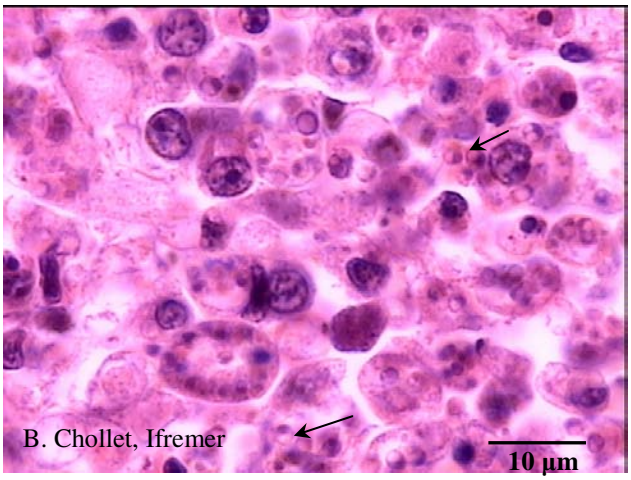


Figure 68. Histological section of the conjunctive tissue around the digestive gland of a *B. ostreae* infected oyster *O. edulis*. Haematoxylin and eosin stain. Arrows correspond to *B. ostreae*.

Haemolymph cell monolayers are prepared by collecting haemolymph from the adductor muscle, and then by cytocentrifugation or cell adhesion onto poly-L-lysine-covered slides. Haemolymph cell monolayers are then fixed and stained with a Hemacolor kit (Merck).

Several authors compared the sensitivity of these 3 light microscopic techniques (O'Neill et al., 1998; Culloty et al., 2003; da Silva and Villalba, 2004) and the use of two different diagnostic techniques reduces false detection of the parasite.

More recently, DNA probe-based techniques have been developed. A PCR reaction for amplifying a rDNA amplicon (528 base pairs (bp) spanning 341 bp of 18S rDNA and 187 bp of ITS1) was proved specific to *B. ostreae* in Europe and North America and more sensitive than histological and tissue imprint techniques (Carnegie et al., 2000). Moreover, phylogenetic analysis of this amplicon confirmed that *B. ostreae* belongs to the phylum Haplosporidia. This amplicon has also been developed into a fluorescent *in situ* hybridization (FISH) assay (Carnegie et al., 2003) specific to *B. ostreae*. Another DNA probe, that amplifies a 300 base pair product, was identified from the same area of the genome by Cochennec et al. (2000). Contrary to the 528 bp fragment identified by Carnegie et al. (2000), this probe also detects other Haplosporidia, *Bonamia* sp. and *Haplosporidium nelsoni*. This probe was successfully used by these authors for *in situ* hybridization and the authors stated the increased power of using this FISH technique in conjunction with routine histopathology, in particular for detecting weak infections. Recently, a real-time TaqMan PCR assay was developed for the detection of *Bonamia* sp. that was comparable to conventional PCR in sensitivity but produced more rapid results with a low risk of sample cross-contamination (Corbeil et al., 2006). This protocol will be optimised to determine the intensity of infection, i.e. to quantify the number of parasites present in the oyster tissues (Arzul, personal communication).

II.2.5. Prophylaxis

Experimental evidence indicates that *B. ostreae* can be transmitted directly between *O. edulis* individuals (Culloty et al., 1999). The occurrence of high mortalities due to *B. ostreae* lead to a change of oyster culture practices such as using suspension culture reducing stocking

densities (Lama and Montes, 1993) or selling oysters at a lower weight (around 40 g), before the peak of *B. ostreae*-induced mortality occurs (Montes et al., 2003).

Despite new management practices, and intensive re-population programmes, the production of *O. edulis* has remained low. Thus there is a need to develop a selective breeding programme for resistance to bonamiasis.

II.3. Selective breeding programme for resistance to bonamiasis (Ifremer)

The main aim of the selective breeding programme for resistance to bonamiasis is the selection of *O. edulis* strains showing a level of tolerance to *B. ostreae* compatible with the re-establishment of the cultivation of this species. The life cycle of *B. ostreae* allows purification of the parasite (Mialhe et al., 1988; Hervio et al., 1995) and thus for experimental infections to be conducted. This technique makes it possible to exert a higher pressure of selection than that which is encountered in the wild, and disease infection can be carried out in a controlled environment. (http://www.ifremer.fr/drvlgp/en/geneticamelioration.htm)

This selective breeding programme was initiated by Ifremer in 1985 and can be divided into two main stages, depending on the way the selection process was achieved.

The first stage of the selective breeding programme consisted of the production of two improved oyster strains (S85 and S89) by mass spawning (Martin et al., 1993; Naciri-Graven et al., 1998). The pressure of selection was applied both by inoculation tests in an experimental hatchery and by field testing in natural conditions. For each selected strain, several generations of selection were obtained, 3 generations for S85 and 2 generations for S89, either produced from parental oysters that had been injected with the parasite (S85i-G2, S89i-G1 and S89i-G2) or not (S85-G1, S85-G2, S85-G3, S89ni-G1 and S89ni-G2) (Figure 69).

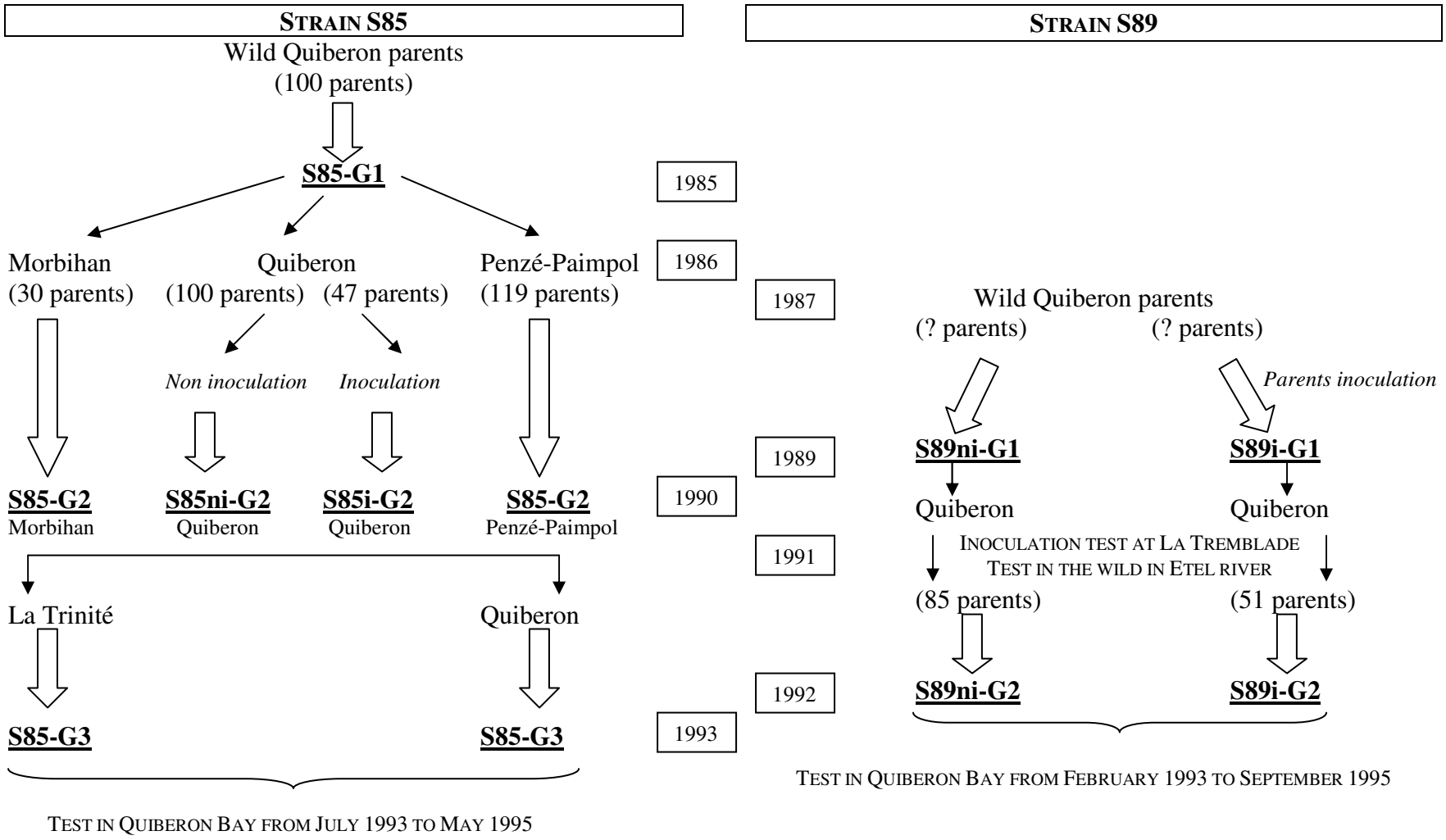


Figure 69. First stage of the programme of selection for *B. ostreae* oysters strains: production of strains S85 and S89 by mass spawning (From Naciri-Graven et al., 1998). ni: issued from non-inoculated parents, i: issued from parents inoculated with the parasite, ?: unknown number of parents.

Several studies reported the success of this first stage of the breeding programme by assessing the relative survival and growth performances of the selected strains S85 and S89 in comparison with control populations that did not undergo artificial selection (Martin et al., 1993; Baud et al., 1997; Naciri-Graven et al., 1998; Naciri-Graven et al., 1999). Both selected strains were associated with a better survival and lower parasite prevalence (Naciri-Graven et al., 1998). At the same time, some studies reported a clear effect of the selection process for *B. ostreae* on growth in the two selected strains, those strains exhibiting better growth performances than the control group (Baud et al., 1997; Naciri-Graven et al., 1999). Two different explanations were discussed by those authors. The first hypothesis stated that difference between the selected and control populations in terms of survival and growth might be due to small differences in infestation rates but also to an indirect selection for a better resistance to experimental conditions (better physiological adaptation of the selected strains to confinement and/or warm environment, because those strains were produced in hatchery). The second hypothesis stated that resistance to bonamiasis was the result of a direct but unintended selection for fast-growing oysters. However, analysis of genetic variability at 5 microsatellite loci revealed that the selected strains, produced by mass spawns from the programme of selection, underwent population bottlenecks. Pedigree reconstruction and estimation of relatedness coefficients revealed a small effective number of breeders, expected to increase inbreeding. Therefore, most indivuals within each strain were half-sibs or full-sibs (Launey et al., 2001) and an important loss of genetic variability in comparison with the original population occurred during the selection process. Those findings lead to a change of strategy of the breeding programme.

The second stage of the selective breeding programme consisted of the production of families on a pair mating scheme, with a within-family based selection process: only individuals that did not die from bonamiasis could be used to generate the next generation. Two batches of production were achieved. Families produced in 1995 (“families 1995”) consisted of intra-strain biparental crosses, meaning that both parents in a cross belonged to the same strain (S85, S89i or S89ni). Then, in order to maximise genetic variability, families produced in 1998 (“families 1998”) issued from inter-strain biparental crosses between S85 and S89 (Figure 70).

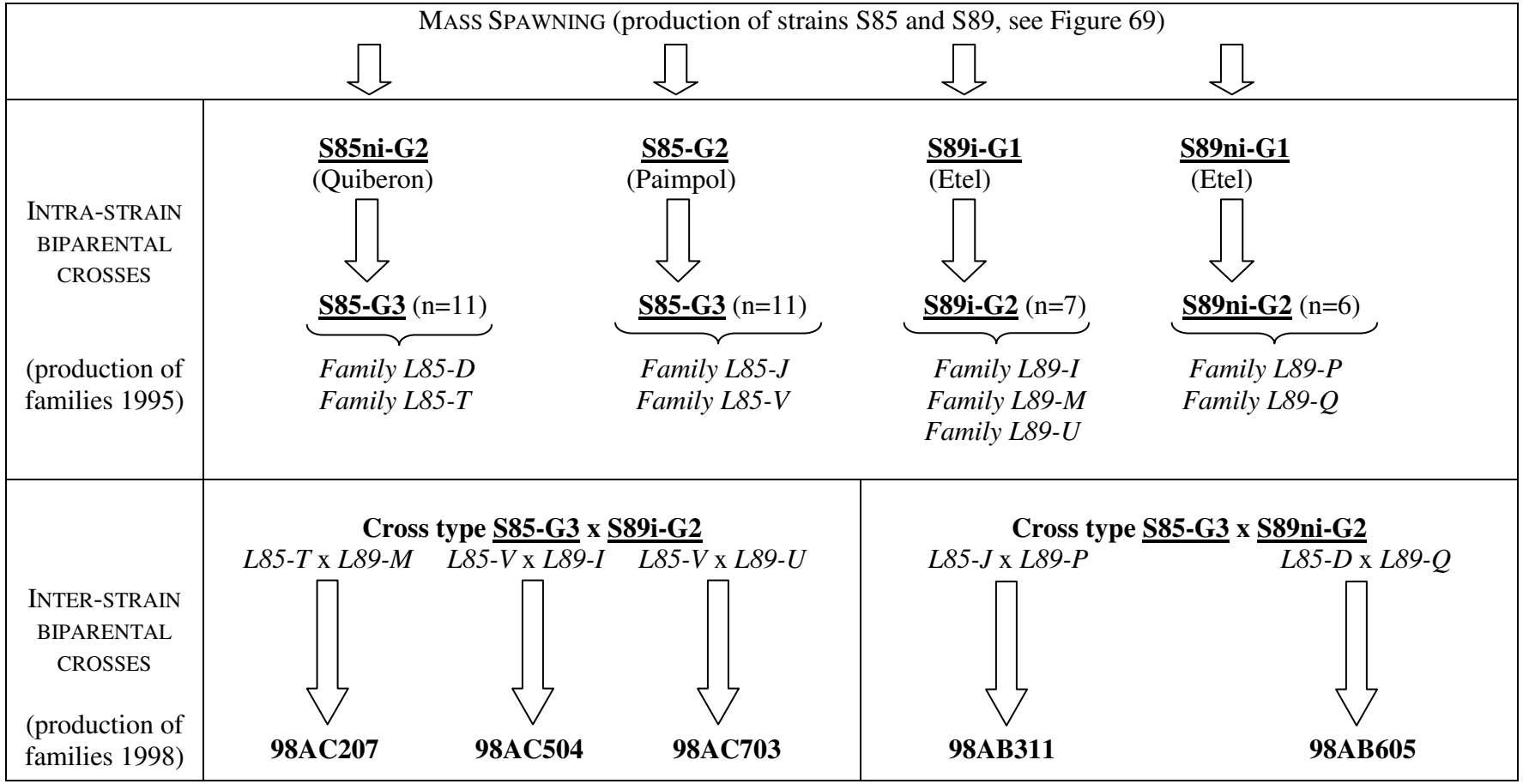


Figure 70. Second stage of selection for *B. ostreae* resistant oysters: production of "families 1995" and "families 1998" by biparental crosses, either intra-strain or inter-strains. See Figure 69 for abbreviations and the history of strains S85ni-G2, S85-G2, S89i-G1 and S89ni-G1.

Bedier et al. (2001) showed that the selected families of *O. edulis* ("families 1998") showed enhanced survival and lower prevalences of the parasite than the control ones in *B. ostreae*-contaminated areas. At the end of the 2 year experimentation, the selected families SS ("families 1998") showed significantly higher overall survival rates (52.3%) than both back-crossed SW (issued from crosses between selected and wild oysters) (40.5%) and control WW (issued from crosses between wild oysters) (2.5%) families. Moreover, *B. ostreae* prevalence was significantly correlated with the survival rates among families. The relative performances of the selected, back-crossed and control families suggested an additive genetic component for the trait (Figure 71).

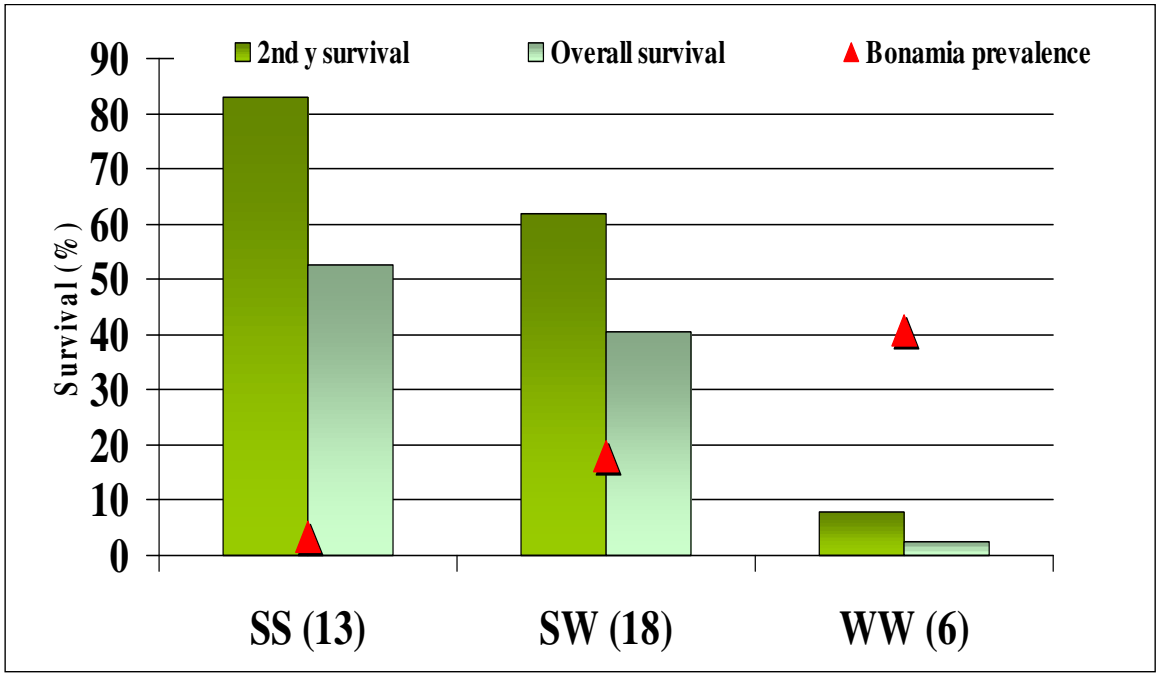


Figure 71. Survival performance in the field of selected (SS), back-crossed (SW) and control families (WW) produced by biparental crosses (Bedier et al., 2001). Families SS correspond to "families 1998" produced by inter-strain biparental crosses between strains S85 and S89. Families SW correspond to crosses between selected and wild oysters and families WW from crosses between wild oysters. Numbers in brackets are the number of families of each type. For each type of family, on the left is shown the 2nd year survival and on the right the overall survival (means are shown). *B. ostreae* prevalence is symbolised by a triangle.

III- MATERIAL AND METHODS

III.1. Production of families F2-S (S=SELECTED)

In 2004, from 20 pairs of oysters, 7 families F1-S (S=SELECTED) were produced by crossing a wild oyster and an oyster from the selected families produced in 1998 (see Figure 70 for the history of these "families 1998"): OE.F1.04.01, OE.F1.04.10, OE.F1.04.15, OE.F1.04.25, OE.F1.04.29 and OE.F1.04.59. In 2005, the second-generation F2-S has been obtained by crossing pairs of fullsibling oysters from the same F1-S family (cross F1-S x F1-S). Fourty pairs of oysters were placed into individual aquariums. Seventeen F2-S families were produced, named OE.F2.05.01, OE.F2.05.02, OE.F2.05.04, OE.F2.05.05, OE.F2.05.06, OE.F2.05.07, OE.F2.05.08, OE.F2.05.09, OE.F2.05.10, OE.F2.05.11, OE.F2.05.12, OE.F2.05.13, OE.F2.05.14, OE.F2.05.15, OE.F2.05.16, OE.F2.05.17 and OE.F2.05.20 (Figure 72).

Two families produced at the experimental hatchery of Ifremer, La Tremblade (France) were put in the "challenge to *Bonamia*" trial experiment, 550 oysters per family:

- an F2-L family derived from a self-fertilisation of an F1 parent: family OE.F2.04.45 (one of the two mapping families used in Chapter 4, see Chapter 4 part III.1.2. for details)
- an F2-S family that was chosen based on DNA polymorphism of microsatellite markers for the grand-parents and F1 parents, and after parentage checking: family OE.F2.05.04 (see above and Figure 72).

III.2. *Bonamia* challenge experiment

This experiment received the help and technical support from the team of pathologists at Ifremer, La Tremblade: Dr Isabelle Arzul, Bruno Chollet, Sylvie Ferrand, Maeva Robert and Benjamin Morga. The success of this experiment depended on the possibility to exert a *Bonamia* pressure on the tested oysters and therefore depended on the purification of *B. ostreae*. The protocol of purification of the parasite was first developed by Mialhe et al. (1988) and has been adapted since then (Figure 73).

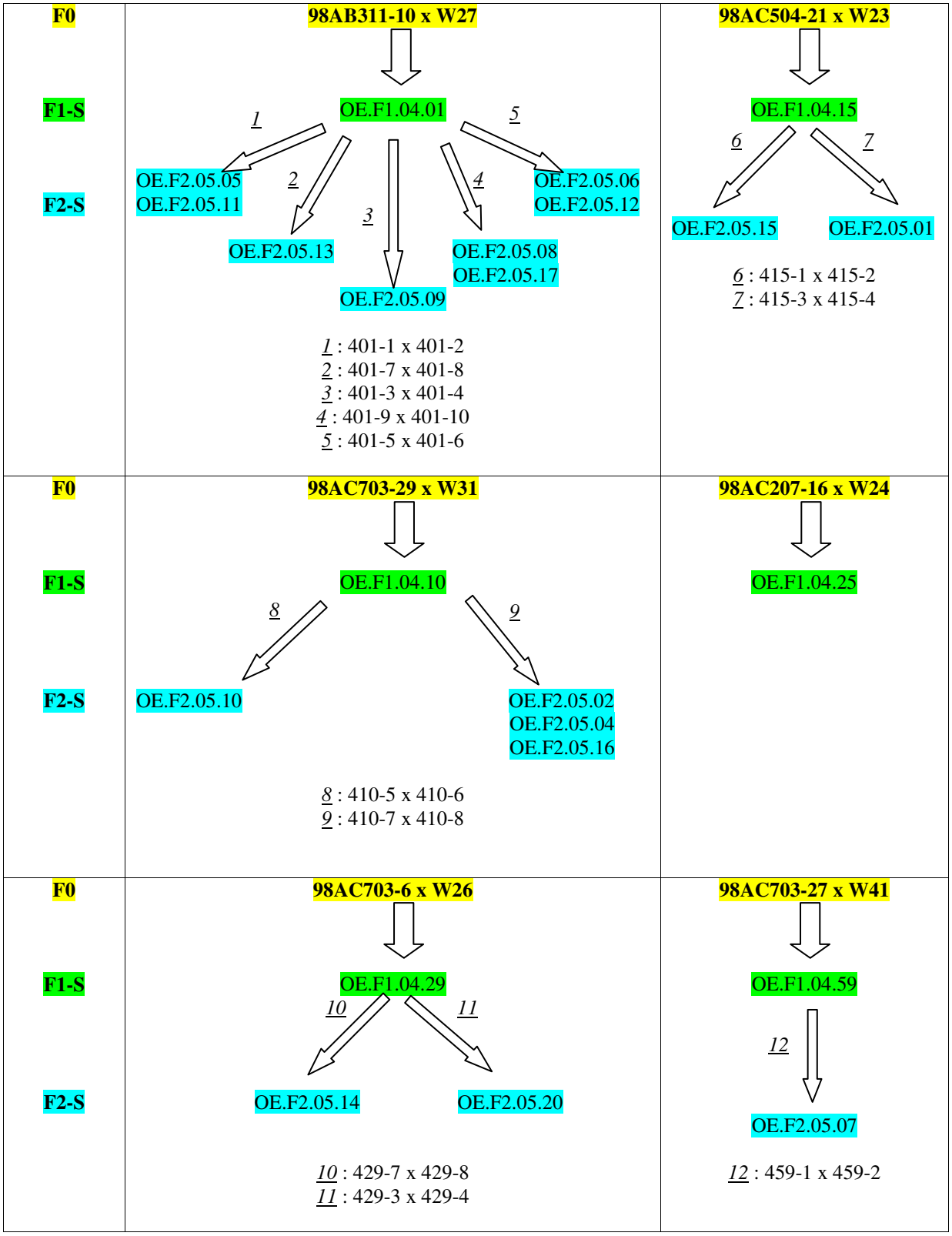


Figure 72. Production in 2005 of 17 F2-S families of *O. edulis* issued from crosses between "families 98" of the programme of selection to bonamiasis (see Figure 70) and wild oysters (W).

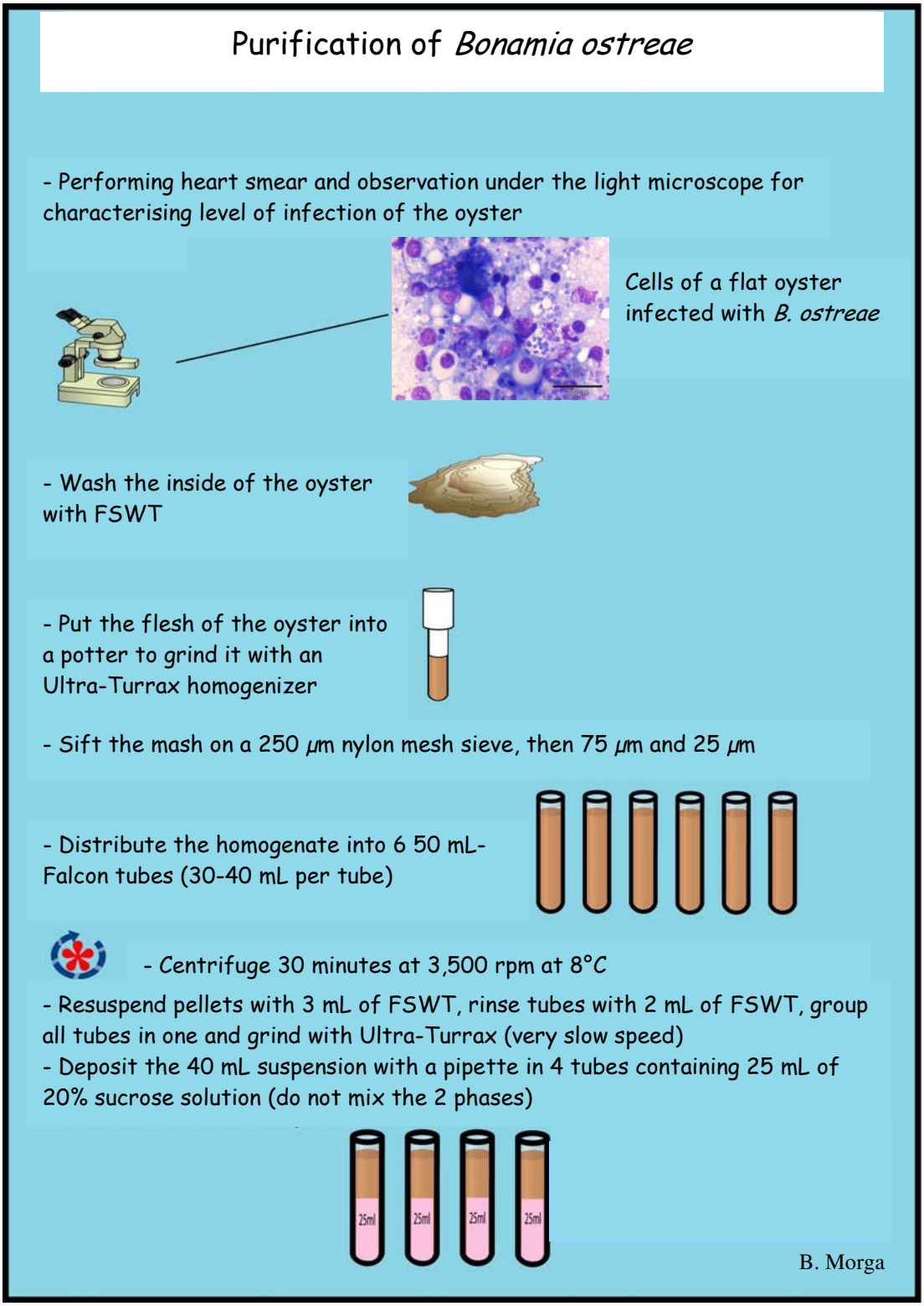


Figure 73. Protocol of purification of the parasite *B. ostreae* (Mialhe et al., 1988; drawings made by Benjamin Morga). FSWT: filtered sea water Tween.

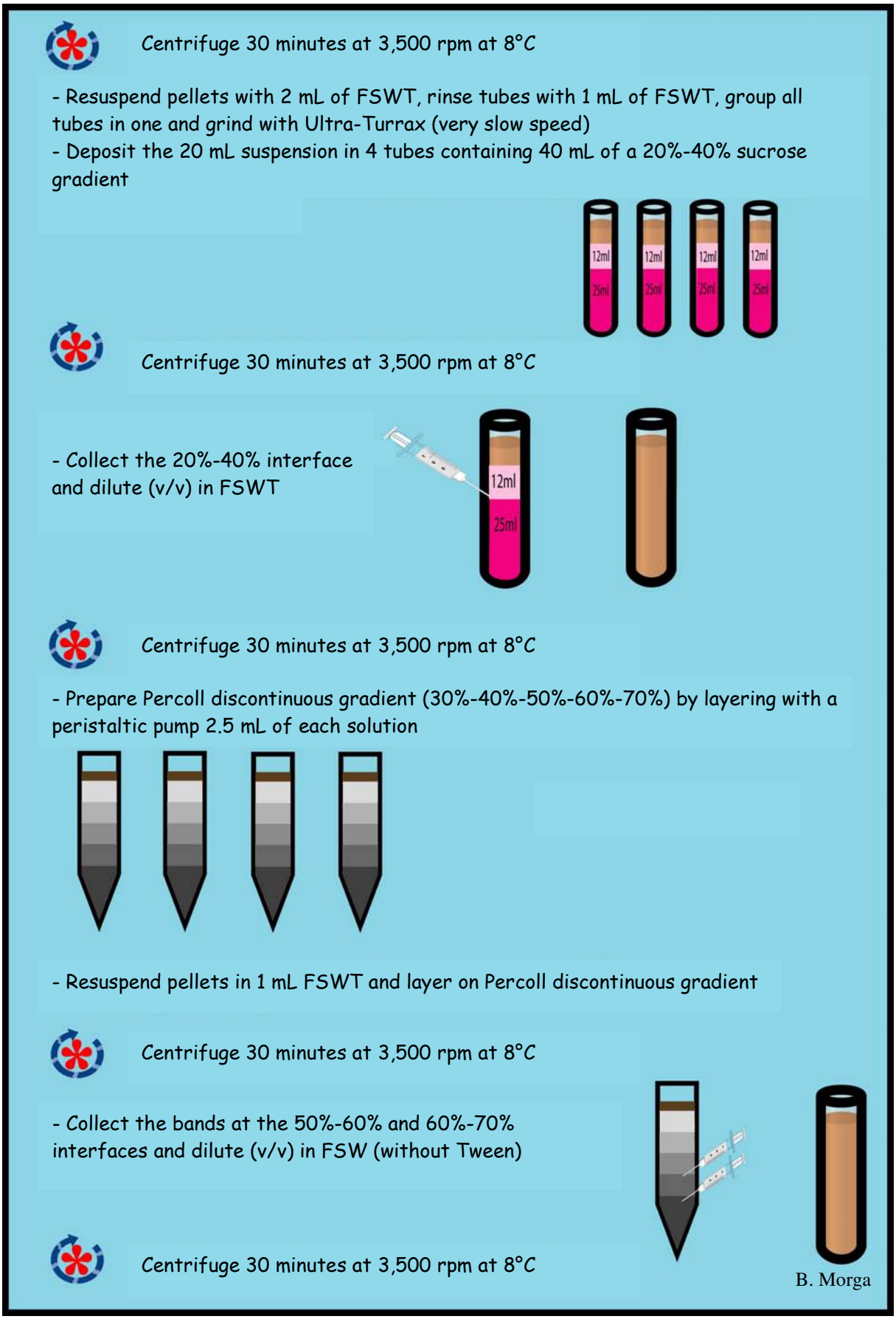
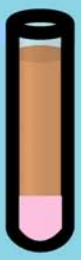


Figure 73- continued

- Empty tubes with a pipette, leave some liquid at the bottom and resuspend pellets in 10 mL of FSW
- Put the suspension on a cushion of 20% sucrose (around 1 mL)



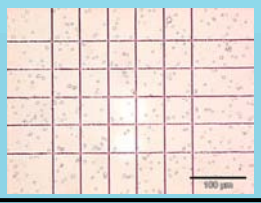
Centrifuge 30 minutes at 3,500 rpm at 8°C

- Empty tubes with a pipette, leave some liquid at the bottom, resuspend pellets in 1 mL of FSW and regroup all fractions in one tube

Centrifuge 30 minutes at 3,500 rpm at 8°C

- Empty tubes with a pipette, dilute the pellet in 500 µL of FSW and rinse the tube with 500 µL of FSW (final volume of 1 mL)

- 1/10 dilution of the purified suspension of *Bonamia* (10 µL *Bonamia*, 90 µL FSW)
- Cell count of the dilution by observation of a Malassez cell at x200 magnification



Count 3 rectangles of 20 squares and take the mean
Take into account the dilution factor

SOLUTIONS TO BE PREPARED:

Filtered sea water Tween (FSWT):
0.2 µm filtered sea water + 1% Tween 80%

Sucrose 20%:
20 g of sucrose in 80 mL FSWT or FSW (for the last cushion)

Sucrose 40%:
40 g of sucrose in 60 mL FSWT

Percoll solution:
For 9 tubes
- 60 mL of Percoll
- 2.46 g of NaCl

Percoll discontinuous gradient:

	Percoll (mL)	SWFT (mL)
30%	6.75	15.75
40%	9	13.5
50%	11.25	11.25
60%	13.5	9
70%	15.75	6.75

B. Morga

Figure 73- end

The experimental design for the "*Bonamia* challenge" experiment consisted of 12 raceways, 6 containing the family OE.F2.04.45 (2 years-old) with 100 oysters per raceway and 5 containing the family OE.F2.05.04 (1 year-old) with 110 oysters per raceway. Each raceway contained 2x2 trays (put on top of each other): the upper tray contained the wild oysters and the lower tray the tested oysters (either family OE.F2.04.45 or family OE.F2.05.04) (Figure 74). Water flow was set up at 150 $l.h^{-1}$ for each raceway. Raceways were cleaned once a week instead of every day in order to increase the opportunity for the parasite to transfer between oysters.

Figure 74. Experimental design of the "*Bonamia* challenge" experiment. Each raceway contained four trays put on top of each other, the 2 upper trays containing the wild oysters (overinfected) and the 2 lower trays the tested oysters (50 per tray for OE.F2.04.45 and 55 per tray for OE.F2.05.04).

The first purification of *B. ostreae* was performed the 16th January 2006 by Bruno Chollet and lead to the collect of 443.10^6 parasites. The cohabitation experiment was started the 17th

January by injecting 1.10^6 parasites into the heart cavity of each wild oyster (from Quiberon bay, Brittany, France). For the 5 raceways containing the family OE.F2.05.04, 17 injected wild oysters and 5 non-injected wild oysters were put in each upper tray (the lower tray containing 55 tested oysters). For the 6 raceways containing the family OE.F2.04.45, 16 injected wild oysters and 6 non-injected wild oysters were put in each upper tray (the lower tray containing 50 tested oysters). Three other purifications of *B. ostreae* performed the 29th March, 7th April and 12th April 2006 by Benjamin Morga and myself lead to 30.10^6, 60.10^6 and 50.10^6parasites and allowed us to inject 8 more wild oysters in total per basket ($0.5.\ 10^6$ parasites). This second batch of injected oysters was added to the upper trays of each raceway to maintain and increase the *Bonamia* pressure.

III.3. Detection of *B. ostreae* in dead oysters

Mortality was checked daily from the 18th January to the 31st July 2006, corresponding to a 6-month trial. The 4 trays were lifted out from the water and were allowed to drain for at least 30 minutes on the side of the raceway before inspecting the oysters. Dead oysters were gaping and were brought to the pathology team to perform a heart smear. Heart smear consisted of taking the ventricle, drying it on paper and then performing several imprints on a slide. After air drying for 5 minutes, the slide was then stained with an Hemacolor kit. The observation was made under the light microscope and the level of infection of the parasite was characterised according to Hervio et al. (1995):

- negative results ($B0^-$) when no parasite was detected after 5 minutes of observation
- low infections ($B0^+$) when 10 or fewer parasites were observed during 5 minutes
- moderate infections ($B0^{++}$) when a few parasites were observed in each microscopic field
- heavy infections ($B0^{+++}$) when several or numerous parasites were observed in each microscopic field

During the week-end, dead oysters were kept individually in foil in the fridge until the Monday morning for the heart smear to be performed. Total shell length (from hinge to outer shell edge) of each dead oyster was measured. The number of the raceway was recorded for each dead oyster, in order to follow daily mortality in each raceway. Cumulative mortality curves were computed to follow the dynamics of mortality as well as the dynamics of appearance of *Bonamia* in each raceway.

The challenge experiment was stopped the 1st August by killing all the surviving oysters of the family OE.F2.05.04. Heart smears were performed for each oyster over 1.5 days by Sylvie Ferrand, Maeva Robert and myself. Observation under the light microscope to determine the level of parasite infection was begun straight away by Bruno Chollet. All observations were made in less than a week.

For the 92 oysters of the family OE.F2.05.04 kept for genetic and QTL mapping, the level of infection was checked by a PCR method, using the 2 pairs of primers specific for *B. ostreae*: pair BO (5'CATTTAATTGGTCGGGCCGC) /BOAS (5'TCTGATCGTCTTCGATCCCCC) (Cochennec et al., 2000) that amplify a portion of the 18S rDNA and SBO (5'GTAATCTTCAACGCGCACCC) /Ra58 (5'CGCATTTCGCTGCGTTCTTC) that amplify the ITS-1 rDNA (ribosomal internal transcribed spacer 1 DNA) (Figure 75). PCR was performed in a 50 µl mix containing about 100 ng of oyster DNA mixed with 5 µl of 10X PCR buffer, 5 µl of 15 µM MgCl2, 5 µl of 2 mM dNTP, 5 µl of each 4 µM primer (BO/BOAS or SBO/Ra58) and 0.5 µl (1 unit) of *Taq* DNA polymerase (Goldstar). Samples were denatured for 5 minutes at 94°C and amplified by 30 cycles of denaturation (1 minute at 94°C), primer annealing (1 minute at 55°C) and elongation (1 minute at 72°C). Complete elongation was ensured by a final extension of 10 minutes at 72°C. Agarose gel (2%) electrophoresis was performed on PCR products alongside to a 100-bp molecular weight standard. Samples containing *B. ostreae* exhibited a 300-bp band with BO/BOAS and a 476-bp band for SBO/Ra58. Only oysters samples for which heart smear and PCR gave the same result were kept for genotyping at molecular markers.

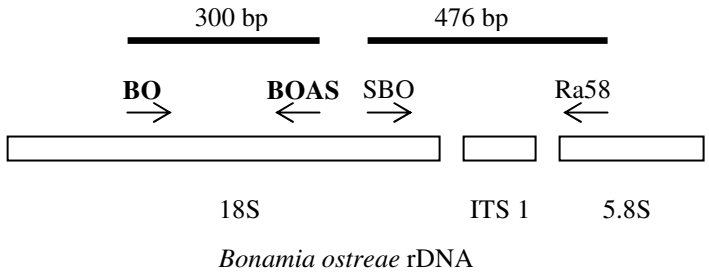


Figure 75. Diagrammatic representation of the *B. ostreae* rDNA with the position of the 2 primer pairs used for PCR detection of the parasite in oyster samples. Primer pair BO/BOAS amplified a portion of the 18S rDNA and lead to an amplicon of 300 bp. Primer pair SBO/Ra58 amplified the ITS-1 rDNA and lead to an amplicon of 476 bp.

III.4. Molecular markers

Samples from the family OE.F2.05.04 consisted of the 2 grand-parents (98AC703-29 and W31), the 2 F1 parents (410-7 and 410-8) of the F1 family OE.F1.04.10 together with 46 F2 progeny that died highly infected with the parasite (diagnosed $B0^{+++}$ after heart smear, and infection confirmed by PCR with primers BO/BOAS and SBO/Ra58) and 46 F2 progeny that survived (parasite could not be found by heart smear or PCR). This three-generation family was genotyped at 20 microsatellites and with 34 different AFLP primer combinations: A1 to A5, A8 to A12, B1 to B5, B8 to B12, E1 to E5, E8 to E12, C1, C5, D1 and D5 (see Chapter 2 for protocols and Table 12). Because the family OE.F2.04.45 proved unsuitable for genetic mapping due to the loss of informativeness resulting from self-fertilisation, only the family OE.F2.05.04 was considered for further QTL mapping.

III.5. QTL analysis

Firstly, a 3-step testing strategy was performed marker by marker to identify potential QTL of bonamiasis resistance or susceptibility, as described in Moen et al. (2004a) for AFLPs. A transmission disequilibrium test (TDT) was applied on the affected offspring (oysters that died heavily infected with the parasite) for all the AFLP markers segregating in that family. The aim was to identify markers that were distorted in the affected oysters, meaning that an allele (band presence) was preferentially transmitted or not transmitted to the affected offspring.

For the AFLPs of type 1:1 (only one of the 2 parents exhibited the band), the TDT was a McNemar test:

$$\chi^2 = \frac{(T - NT)^2}{(T + NT)}$$

where T is the number of times an allele (band presence) is transmitted from a heterozygous parent (Aa) to an affected offspring; and NT the number of times this allele is not transmitted.

For the AFLPs of type 3:1 (both parents exhibited the band), the TDT was based on a likelihood ratio test (LRT):

$$LRT = -2\ln\frac{L(v_s = 0.5, v_d = 0.5)}{L(v_s, v_d)}$$

where $L(v_s, v_d) = K_1((1 - v_s)(1 - v_d))^{n_a}(1 - (1 - v_s)(1 - v_d))^{n_p}$

with K_1 being a constant, n_a being the number of affected offspring having the band-absent marker phenotype and n_p the number of affected offspring having the band-present marker phenotype, v_s being the probability of transmission of the sire "band-present" allele to its offspring and v_d the probability of transmission of the dam "band-present" allele to its offspring. Therefore $(1\text{-}v_s)(1\text{-}v_d)$ corresponds to the probability of having a null homozygote offspring (genotype *aa*) or "band-absent" marker phenotype. $L(0.5,0.5)$ represents therefore the probability that both the sire and dam have a probability of 0.5 of transmitting the band present allele (as expected); $L(v_s,v_d)$ is the observed joint probability of transmission of the sire and dam. For example, consider an AFLP marker for which 20 affected offspring exhibit the band, and one affected offspring do not exhibit the band; the *LRT* will be:

$$LRT = -2\ln\frac{K_1(0.5x0.5)^1(1-(0.5x0.5))^{20}}{K_1(1/21)^1(1-(1/21))^{20}} = 6.24$$

McNemar test and LRT approximately follow a chi-square distribution with 1 degree of freedom. The following decision rule will be applied:

v=1	**Test**	**Interpretation**	**Action**
$\chi^2 < 3.841$	NS	No distortion in the transmission	Marker discarded
$3.841 < \chi^2 < 6.635$	p<0.05	Significant TDT; there is a distortion in the transmission from the parents to the affected offspring	Marker kept for the second step of the multi-step testing strategy
$6.635 < \chi^2 < 10.828$	p<0.01		
$\chi^2 > 10.828$	p<0.001		

The second test consisted of a Mendelian segregation test (MST) that was applied on markers significant after the TDT. This test was performed on the whole dataset, i.e. the 46 affected offspring (that died heavily affected with the parasite) and the 46 non-affected offspring (that survived the trial experiment and were proved parasite-free). The test was performed as explained in the part III.3. of Chapter 4. Only markers that were significant after the TDT and not significant after the MST were kept for the survival analysis. This was to ensure removal of markers that were distorted for reasons other than the linkage to potential genes of resistance/susceptibility to the disease. Indeed, a marker that was distorted when considering the affected offspring only and no more distorted when the non-affected offspring were included means that the distortion was acting in opposite directions in the affected and non-affected offspring. For example, the band-absent phenotype could be more frequent than expected in the

affected offspring and the band-present phenotype more frequent than expected in the non-affected offspring: the two distortions will cancel each other and the marker will segregate according to Mendelian rules in the whole dataset (affected + non-affected).

The third test was the survival analysis itself. This test uses more information than the TDT does, because the variable used is the number of days an offspring survived and both affected and non-affected offspring are taken into account. Survival of two groups of offspring was compared: offspring with the band-present marker phenotype and offspring with the band-absent marker phenotype. This analysis was performed among progeny of a single fullsib family sharing the same environment. Hypothesis testing was carried out using a log-rank test:

$$LRANK = \frac{(O_a - E_a)^2}{E_a} + \frac{(O_p - E_p)^2}{E_p}$$

where O_a is the total number of affected offspring in the band-absent group, O_p the number of affected offspring in the band-present group, E_a and E_p are the respective expected counts. The expected count E_a was computed as followed:

$$E_a = \sum_{i=1}^{t} E_{ai} \text{ with } E_{ai} = d_i \left(\frac{r_{ai}}{r_i} \right)$$

where d_i is the number of offspring that died during day i, r_i the total number of offspring still alive at the beginning of day i and r_{ai} the number of offspring still alive in the band-absent group at the beginning of day i. Similarly, the expected count E_p was computed as:

$$E_p = \sum_{i=1}^{t} E_{pi} \text{ with } E_{pi} = d_i \left(\frac{r_{pi}}{r_i} \right)$$

where d_i is the number of offspring that died during day i, r_i the total number of offspring still alive at the beginning of day i and r_{pi} the number of offspring still alive in the band-present group at the beginning of day i.

LRANK was distributed as a chi-square distribution with one degree of freedom. Moreover, Kaplan-Meier survival curves were constructed for both groups by computing the cumulative proportion of offspring within the group (band-present or band-absent) still alive at any time:

$S_a(t_i) = S_a(t_{i-1}) \frac{(r_{ai} - d_{ai})}{r_{ai}}$ for the band-absent group

and $S_p(t_i) = S_p(t_{i-1})\frac{(r_{pi} - d_{pi})}{r_{pi}}$ for the band-present group

The difference between the survival curves for the band-present and band-absent group was estimated by the hazard ratio as stated in Moen et al. (2004a):

$$h = \frac{(O_p / E_p)}{(O_a / E_a)}$$

Having robustly identified potential QTLs for resistance or susceptibility to bonamiasis with the three-step testing strategy of Moen et al. (2004a) (TDT, MST and survival analysis), a genetic linkage map was built for the family OE.F2.05.04 with CriMap software. All microsatellites, all AFLPs of type 1:1 and the Mendelian 3:1 AFLPs were considered for linkage analysis. Two parental maps were constructed, one for the F1 parent 410-7 and one for 410-8.

Finally, a QTL mapping approach was performed with the QTL express software (Seaton et al., 2002) (http://qtl.cap.ed.ac.uk). Three input files were prepared:

- a genotype file containing the number of markers, list of names of the markers, names given to each generation (P1, P2, F1, F2), names given to males and females (0 in our case because sex of individuals was unknown), symbol for missing genotype data (0) and genotypic data for each individual (one line per individual with individual ID, name of parents, sex, name of generation and genotypes)
- a map file containing the number of chromosomes, the interval for calculation of genotype probabilities and coefficients (1 cM) and then 2 lines for each chromosome, the first line containing the name of the chromosome and the number of markers, the second line containing markers in order separated by distances between them in cM
- a phenotype file containing the number of fixed effects, the number of covariates and the number of traits; names of each fixed effect, covariate and trait; code for a missing trait value (-999); and then one line per individual containing the individual ID and values for the fixed effects, covariates and traits.

In our study, one fixed effect was tested: raceway in which the oysters were kept (5 different raceways: 8, 11, 12, 13 and 14) and one covariate: total length at the time of death (in cm). The trait analysed was binary: the oysters that survived the challenge experiment were coded "0" and the ones that died during the challenge experiment were coded "1".

The module "Large Single Full-Sib Family Analysis (Tree)" was chosen because it is designed for the case of a family with two parents that are not assumed to come from a cross between two distinct and different genetic lines (as they would if we were analysing an F2 cross between two lines). With this module, three estimates can be obtained, the paternal estimate, the maternal estimate and the interaction estimate. Indeed, the analysis makes a comparison between the two gametes carried by the male parent (the paternal component) and the two gametes carried by the female parent (the maternal component). The grandparent information is used to help identify the two parental gametes - so the first paternal gamete comes from grandparent 1, the second paternal gamete comes from grandparent 2 and similarly for the maternal comparison. If the paternal component is significant it suggests that the two different paternal gametes carry alleles of different effect for the QTL. A similar conclusion for the maternal component can be made. In this analysis we do not assume that both the paternal and maternal comparison will be significant at the same time and if they are significant then we do not assume that they will be in the same direction. The last component – the interaction effect - allows for an interaction between the maternal and paternal components. The most likely reason for this is dominance. For example, if the gene for resistance is recessive and both father and mother are heterozygous for the QTL, then we may see on average a paternal effect and on average a maternal effect, but we see a much greater effect when we have both the resistance alleles from both parents, hence there is an interaction (Haley, personal communication). To estimate these three components, four different models are available in QTL express: pat (to estimate the paternal component only), mat (to estimate the maternal component only), pat + mat (to estimate both the paternal and maternal components) and pat + mat + int (to estimate the interaction component).

The model pat + mat assumes that the map is the same in the two parents, so that a consensus map could be built. In our study, only a few microsatellites were used (that could serve as anchor loci between the 2 parental maps) and most molecular markers were AFLPs. Therefore,

2 parental maps were built, one for the parent 410_7 and another one for the parent 410_8 and the two maps could not be integrated because some markers were in common (microsatellites and a few AFLPs) but the rest were informative only in one parent or the other. Therefore, separate analyses were performed for the two parents:

- for the parent 410_7, 3 input files were used: the genotype file contained the genotypic data for the 2 grand-parents (703-29 and W31), the 2 parents (410_7 and 410_8) and the 92 F2 progeny for the molecular markers that were mapped into the 410_7 map; the map file contained the linkage groups for the 410_7 map; and the phenotype file. The "pat" model was used to fit the paternal component and to find QTLs.
- for the parent 410_8, 3 input files were used: the genotype file contained the genotypic data for the markers that were mapped into the 410_8 map; the map file contained the linkage groups for the 410_8 map; and the phenotype file. The "mat" model was used to fit the maternal component and to find QTLs.

The paternal component relating to 410_7 and the maternal component to 410_8 were arbitrary (sex unknown in our dataset) and corresponded to the order in which the parents were entered in the genotype input file after the individual ID.

The finding of QTLs used a regression interval mapping approach (Haley and Knott, 1992; Martinez and Curnow, 1992). The analysis was based on least squares methods for performing flanking marker analyses for QTL detection. It uses information from 2 linked flanking markers, and the presence of a QTL in this interval is tested. The test consists of fitting the one-QTL linear model to find the parameter estimates that minimise the residual sums of squares (RSS); if the presence of a QTL in that interval is accepted, the position associated to the minimum RSS gives the most likely position of a QTL as well as the best estimates of its effects. Chromosome-wide significance threshold (which takes into account multiple testing on a specific chromosome) was estimated after performing 1000 permutations according to Churchill and Doerge (1994), by permuting the array of coefficients for an individual over phenotypes.

IV- RESULTS

IV.1. *Bonamia* challenge experiment: daily check-up of mortality

Total cumulative mortality and cumulative mortality of oysters that died infected with the parasite are shown in Figure 76 for the first batch of wild oysters (that were injected with the parasite) and for the 2 tested families, OE.F2.04.45 and OE.F2.05.04. Mortality in the tested families began in May (4 months after the beginning of the trial). Mortality of oysters that were *Bonamia*-positive occurred in June and July. In total, 52 oysters of the family OE.F2.04.45 died (9.4%). Smears could not be performed for 11.5% of them because of the degree of tissue decay of the oyster; 28.85% were diagnosed $B0^-$, 28.85% $B0^+$, 23.1% $B0^{++}$ and only 7.7% were $B0^{+++}$. For the family OE.F2.05.04, 105 oysters in total died over the 6 month period (19.1%). Smears could not be performed for 9.5% of them; 9.5% were $B0^-$, 16.2% $B0^+$, 22.9% $B0^{++}$ and 41.9% were $B0^{+++}$. The mortality was significantly higher in the family OE.F2.05.04 (χ^2=24.34, p<0.001, 4 d.f.) as well as the level of infection of the parasite found in heart smear (χ^2=20.87, p<0.001, 1 d.f.).

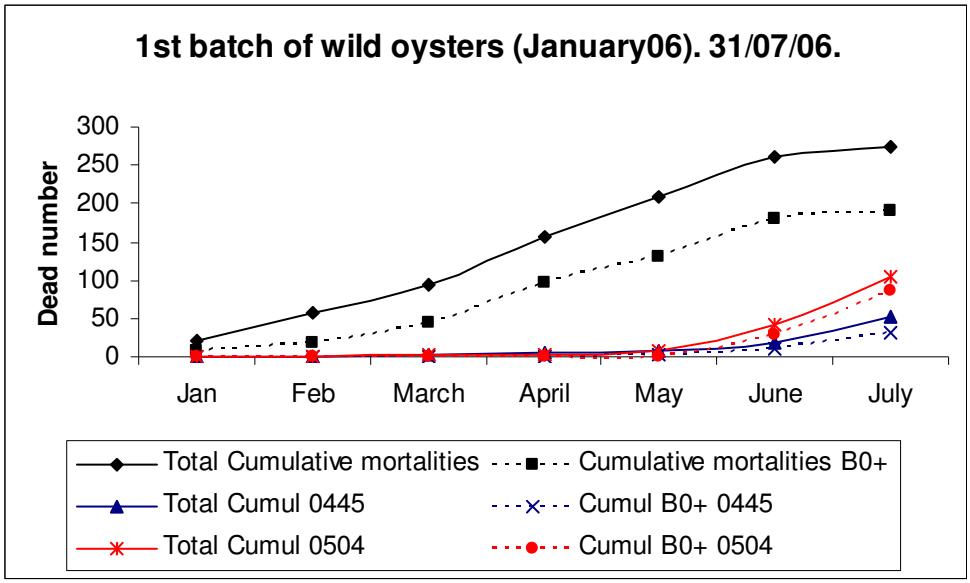


Figure 76. Cumulative mortalities for the 6 month trial of the "*Bonamia* challenge" experiment. The two upper curves (diamond and square) represent the cumulative total mortality and the cumulative mortality of oysters infected with the parasite for the wild oysters (upper basket). The 4 lower curves represent the total mortality and *Bonamia*-infected oysters mortality for the family OE.F2.05.04 (star and round) and the family OE.F2.04.45 (triangle and cross). Data were pooled across all raceways.

Cumulative mortalities for the family OE.F2.04.45 are shown for the 6 raceways individually in Figure 77. All raceways seem to present the same patterns, with the first mortalities occurring in June and July in the tested oysters. Only the raceway 6 showed no $B0^+$ mortality: only one tray of 50 tested oysters was present in that raceway (instead of 2 trays of 50 oysters each): the infective pressure of the parasite was probably reduced because half as many wild oysters were present in this raceway (only one upper tray instead of two).

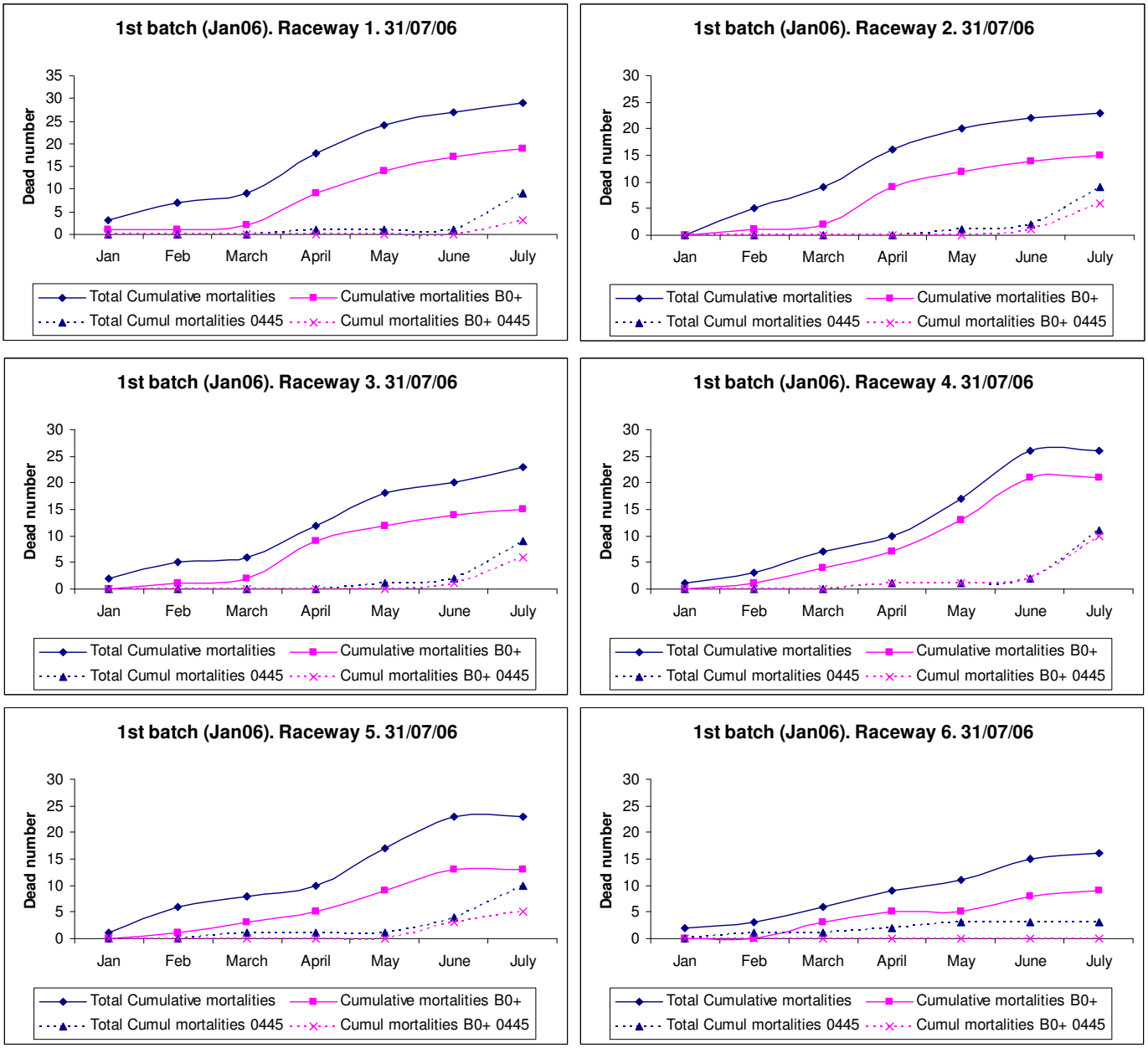


Figure 77. Cumulative mortalities for the 6 month trial of the "*Bonamia* challenge" experiment for the family OE.F2.04.45, per raceway. The two upper curves correspond to the total cumulative mortality and $B0^+$ mortality for the first batch of wild oysters; and the two lower curves for the family OE.F2.04.45.

In the same way, cumulative mortalities were constructed for the family OE.F2.05.04 (Figure 78). All 5 raceways presented a similar dynamics of mortality, with the first mortalities appearing in May but most mortalities occurring in June and July and a parallel evolution of the total cumulative mortality and the $B0^+$ mortality. Moreover, no significant difference in the number of dead oysters (χ^2=4.23, p=0.376, 4 d.f.) or level of infection in the dead oysters (χ^2=14.59, p=0.265, 12 d.f.) was found between the 5 raceways: data were therefore pooled.

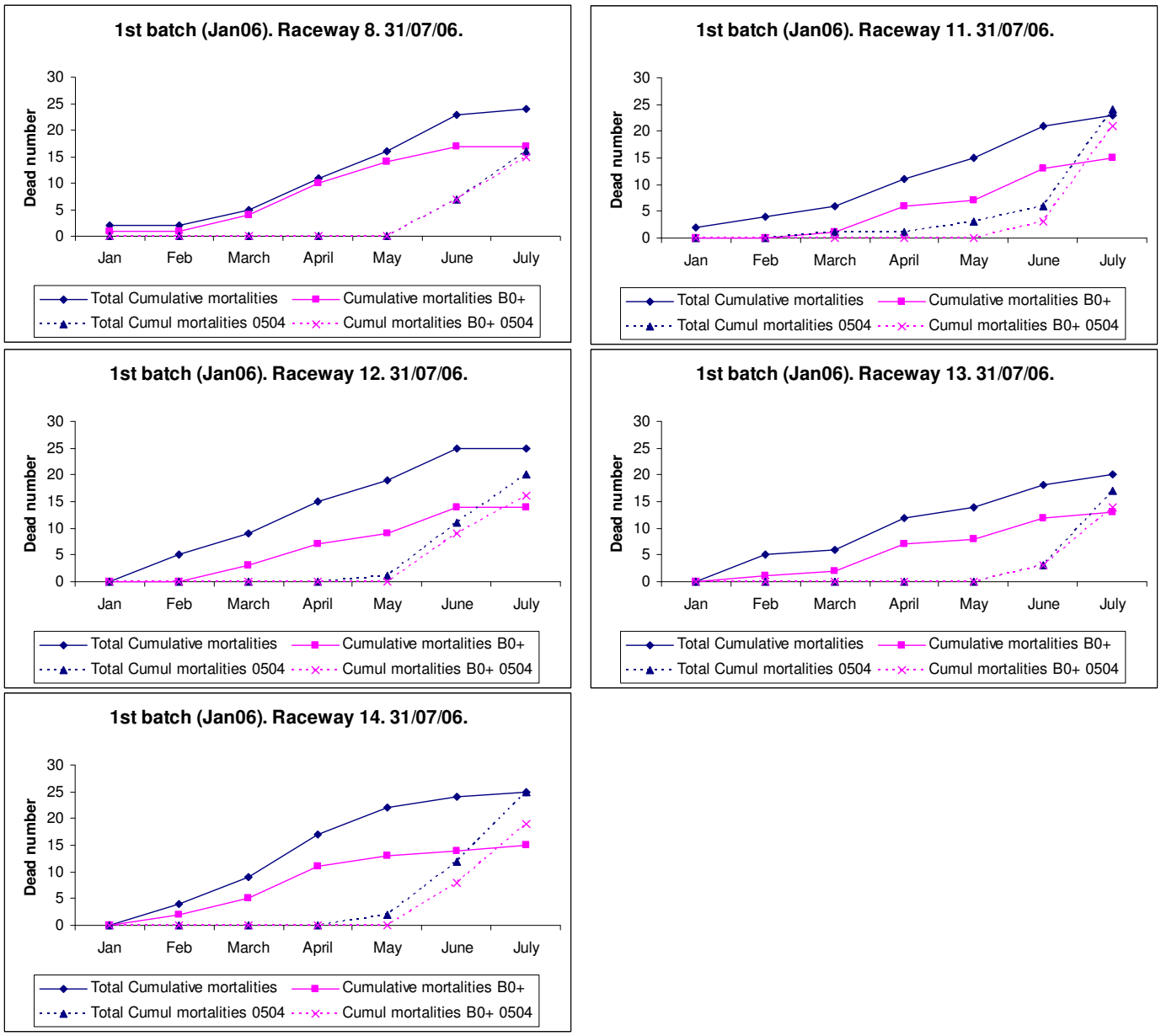


Figure 78. Cumulative mortalities for the 6 month trial of the "*Bonamia* challenge" experiment for the family OE.F2.05.04, per raceway. The two upper curves correspond to the total cumulative mortality and $B0^+$ mortality for the first batch of wild oysters; and the two lower curves for the family OE.F2.05.04.

All the surviving oysters of the family OE.F2.05.04 were opened the 1st August 2006 (after a 6 month trial) and a heart smear performed on them. Overall, 444 oysters were analysed, 94 in raceway 8, 85 in raceway 11, 87 in raceway 12, 95 in raceway 13 and 83 in raceway 14. The vast majority of them were $B0^-$ (86.2% for raceway 8, 89.4% for raceway 11, 93.1% for raceway 12, 87.4% for raceway 13 and 88.0% for raceway 14) and almost no $B0^{+++}$ oysters were found (0% in raceways 8, 11, 12 and 13; and only 1.2% in raceway 14) (Table 29). This means that the experiment was not stopped too early, and that the mean peak of mortality was passed. No significant difference between raceways was found for the heart smear results achieved on the surviving oysters (χ^2=3.92, p=0.864, 8 d.f. after pooling $B0^{++}$ and $B0^{+++}$; χ^2=2.55, p=0.636, 4 d.f. after pooling $B0^+$, $B0^{++}$ and $B0^{+++}$). Therefore, data were pooled. Overall, 88.7% of the surviving oysters were not infected with the parasite ($B0^-$), 8.4% were slightly infected ($B0^+$), 2.7% were moderately infected ($B0^{++}$) and only 0.2% were heavily infected ($B0^{+++}$).

Table 29. Heart smear results on the 444 surviving oysters of the family OE.F2.05.04

Raceway	Total killed oysters	Smear result	No. oysters	Percentage (%)
Raceway 8	94	$B0^-$	81	86.2
		$B0^+$	11	11.7
		$B0^{++}$	2	2.1
		$B0^{+++}$	0	0
Raceway 11	85	$B0^-$	76	89.4
		$B0^+$	6	7.1
		$B0^{++}$	3	3.5
		$B0^{+++}$	0	0
Raceway 12	87	$B0^-$	81	93.1
		$B0^+$	4	4.6
		$B0^{++}$	2	2.3
		$B0^{+++}$	0	0
Raceway 13	95	$B0^-$	83	87.4
		$B0^+$	9	9.5
		$B0^{++}$	3	3.1
		$B0^{+++}$	0	0
Raceway 14	83	$B0^-$	73	88.0
		$B0^+$	7	8.4
		$B0^{++}$	2	2.4
		$B0^{+++}$	1	1.2
Overall	444	$B0^-$	394	88.7
		$B0^+$	37	8.4
		$B0^{++}$	12	2.7
		$B0^{+++}$	1	0.2

The difference in the level of infection between the 2 groups of oysters (dead, surviving) was striking: 80% against 11% of *Bonamia*-positive in the dead or surviving groups respectively. For further genetic analysis, 92 oysters were randomly chosen between the 5 raceways, 46 oysters that died heavily infected with the parasite ($B0^{+++}$) and 46 surviving oysters in which no parasite could be found after heart smear. The 5 raceways were equally represented. The status of infection to the parasite was checked by PCR (primer BO/BOAS): all the $B0^{+++}$ oysters after heart smear showed a very clear band of 300 bp in a 2% agarose gel. On the contrary, the 46 $B0^{-}$ oysters after heart smear were lacking that band. Confirmation was obtained with the primer pair SBO/Ra58. These 92 F2 progeny, the 2 grand-parents 98AC703-29 and W31 and the 2 parents 410-7 and 410-8 were genotyped for 20 microsatellites and 34 AFLP primer pairs (Figure 79).

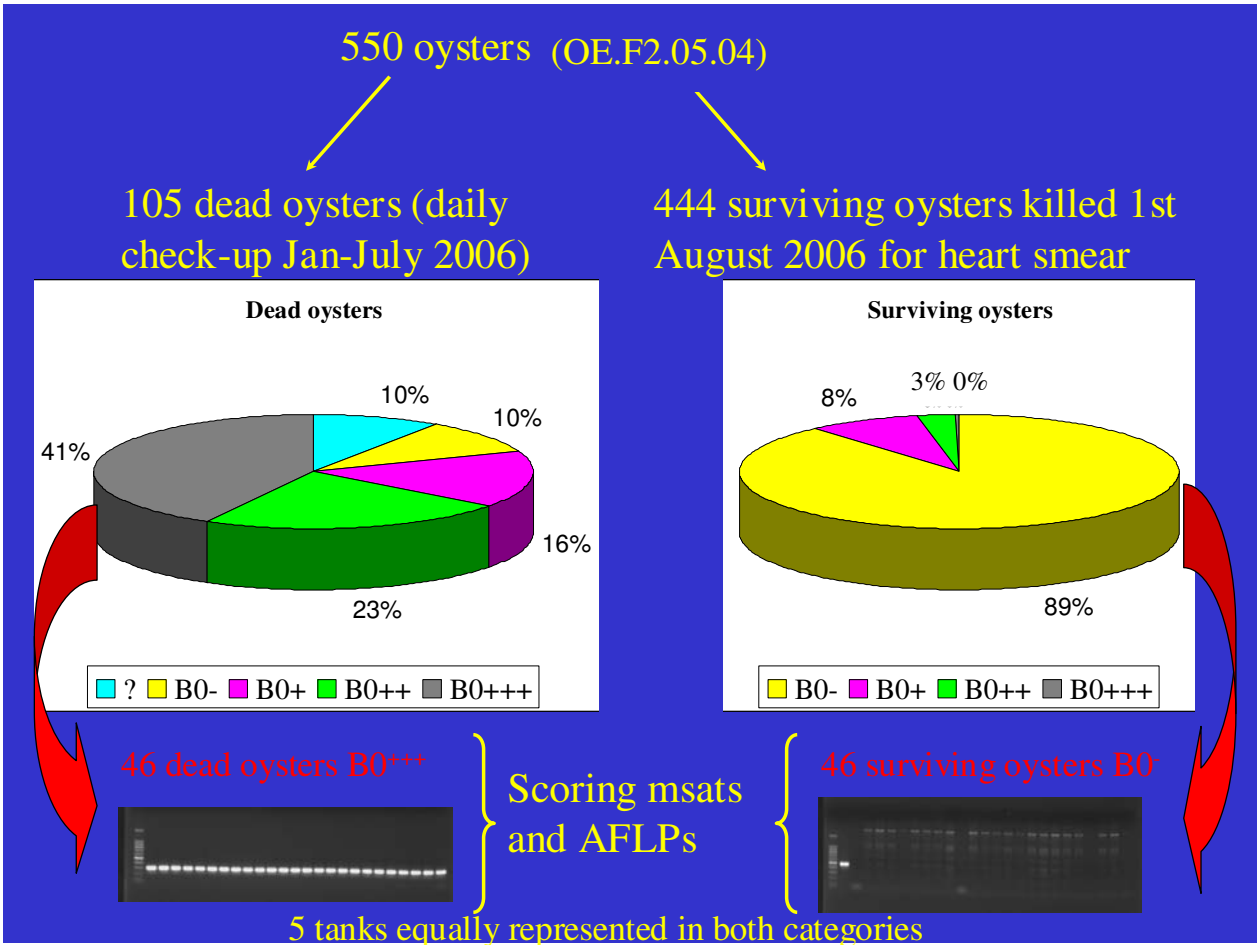


Figure 79. Comparison of level of infection of the parasite (after heart smear) between the 2 groups of oysters of the family OE.F2.05.04, the 105 that died during the 6 month trial and the 444 that survived the challenge experiment. 46 oysters in each group were randomly chosen in the 5 raceways after PCR confirmation of the *Bonamia* status (primers BO/BOAS and SBO/Ra58): molecular weight standard and the positive control (sample infected with the parasite) are shown on the left of the gel.

IV.2. Survival analysis and the search for potential QTLs of resistance/susceptibility to bonamiasis

The 34 AFLP primer pairs produced 309 markers, 201 AFLPs of type 1:1 and 108 of type 3:1. After the Transmission Disequilibrium Test (TDT), 144 markers were significant and kept for the Mendelian Segregation Test (MST): 83 of type 1:1 (22 with $p<0.05$; 20 with $p<0.01$; 41 with $p<0.001$) and 61 of type 3:1 (16 with $p<0.05$; 18 with $p<0.01$; 27 with $p<0.001$). Of the 144 markers significant after the TDT, only 26 were not significant after the MST: 5 markers segregated through the parent 410-7 (1:1 type), 10 through the parent 410-8 (1:1 type) and 11 through both parents (3:1 type) (Table 30).

Table 30. TDT and MST for the 26 markers that were kept for survival analysis (significant after TDT among susceptible progeny and non significant after MST in the whole progeny). present: band-present phenotype (Aa or A?), absent: band-absent phenotype (aa); TDT: transmission disequilibrium test; MST: Mendelian segregation test; *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$.

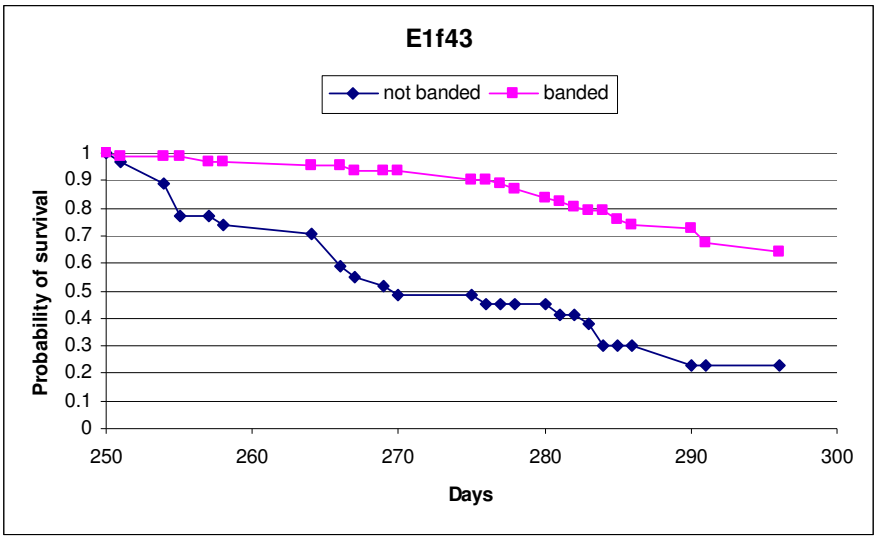
	Parental genotypes		No. of susceptible offspring (dead)		No. of resistant offspring (surviving)			
Marker	410-7	410-8	present	absent	present	absent	TDT	MST
A1f150	aa	Aa	33	12	21	25	9.8**	3.1
E1f43	Aa	Aa	25	21	38	8	9.2**	2.1
A3f73	aa	Aa	12	34	25	21	10.5**	3.5
A3f165	Aa	Aa	42	4	35	11	8.1**	3.7
E3f169	aa	Aa	10	36	27	19	14.7***	3.5
E3f255	Aa	Aa	25	21	38	8	9.2**	2.1
E4f291	Aa	aa	16	30	23	23	4.3*	2.1
A5f225	aa	Aa	16	30	21	25	4.3*	3.5
E5f126	Aa	Aa	40	6	33	13	4.0*	0.9
E5f157	aa	Aa	13	33	24	22	8.7**	3.5
B8f234	Aa	Aa	43	3	32	14	10.9***	2.1
E9f147	Aa	Aa	40	6	37	9	4.0*	3.7
E9f368	aa	Aa	9	37	28	18	17.0***	3.5
E9f371	Aa	aa	31	15	22	24	5.6*	2.1
A12f288	Aa	Aa	43	3	33	13	10.9***	2.8
A12f429	Aa	Aa	25	21	38	8	9.2**	2.1
B12f52	aa	Aa	16	30	21	25	4.3*	3.5
B12f243	aa	Aa	37	9	17	29	17.0***	2.8
B12f478	Aa	aa	31	15	22	24	5.6*	2.1
C1f99	aa	Aa	37	9	18	28	17.0***	3.5
D1f129	Aa	Aa	42	4	33	13	8.1**	2.1
D1f162	Aa	aa	31	15	23	23	5.6*	2.8
D1f203	Aa	Aa	22	24	40	6	15.5***	2.8
D1f328	aa	Aa	12	34	30	16	10.5**	0.7
C5f112	Aa	aa	31	15	24	22	5.6*	3.5
D5f203	Aa	Aa	40	6	37	9	4.0*	3.7

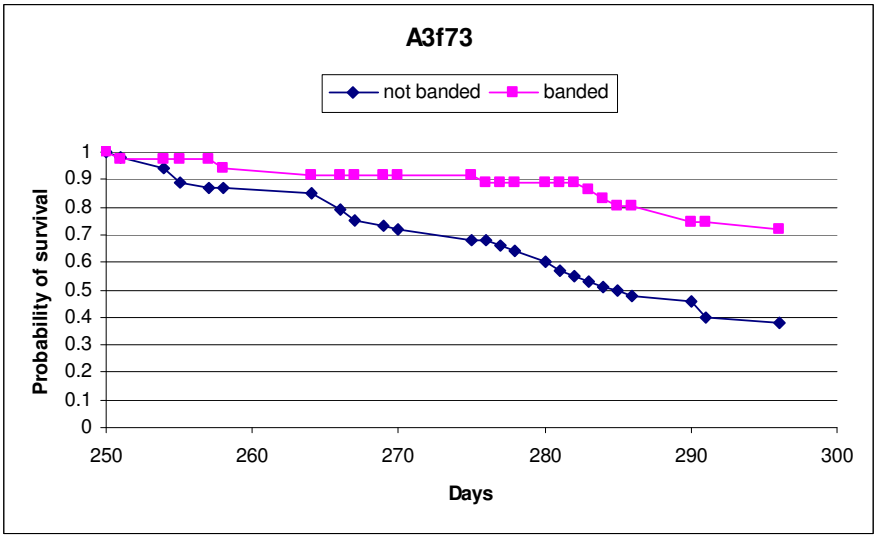
Of the 26 markers kept for survival analysis, 15 were significant (4 with $p<0.05$, 2 with $p<0.01$ and 9 with $p<0.001$). The band-present marker phenotype corresponded to a resistance allele in 9 cases and to a susceptible allele in 6 cases. Hazard ratios were in the range 0.24-0.45: inheriting the band-present allele reduced (or increased) the mortality by ~ 24% to 45% for a resistant (or susceptible) marker (Table 31).

Table 31. Survival analysis results for the family OE.F2.05.04, hazard ratio and LRANK. R: resistance marker, S: susceptible marker; O_a: total number of dead offspring in the band-absent marker phenotype and E_a its relative expected count; O_p: total number of dead offspring in the band-present marker phenotype and E_p its relative expected count. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$.

Marker	Origin	R or S marker	O_a	E_a	O_p	E_p	Hazard ratio (h)	LRANK
A1f150	W31, 410-8	S	12	21.8	33	24.2	0.40	7.7**
E1f43	703-29, 410-7, 410-8	R	21	10.0	25	36.0	0.33	15.5***
A3f73	703-29, 410-8	R	34	24.2	12	21.8	0.39	8.4**
E3f169	W31, 410-8	R	36	24.1	10	21.9	0.31	12.3***
E3f255	703-29, 410-7, 410-8	R	21	9.5	25	36.5	0.31	17.6***
E5f157	703-29, 410-8	R	33	24.5	13	21.5	0.45	6.4*
B8f234	703-29, 410-7, 410-8	S	3	10.1	43	35.9	0.25	6.4*
E9f368	W31, 410-8	R	37	23.6	9	22.4	0.26	15.6***
A12f288	703-29, 410-7, 410-8	S	3	9.5	43	36.5	0.27	5.5*
A12f429	703-29, 410-7, 410-8	R	21	10.0	25	36.0	0.33	15.5***
B12f243	703-29, 410-8	S	9	23.1	37	22.9	0.24	17.3***
C1f99	703-29, 410-8	S	9	22.4	37	23.6	0.26	15.6***
D1f129	703-29, 410-7, 410-8	S	4	9.7	42	36.3	0.36	4.3*
D1f203	W31, 410-7, 410-8	R	24	12.6	22	33.4	0.34	14.3***
D1f328	W31, 410-8	R	34	21.5	12	24.5	0.31	13.7***

Kaplan-Meier survival curves were constructed for the 9 AFLP markers for which the band-present phenotype corresponded to a resistant allele. In those cases, the band-present phenotype was associated with better survival of between 26% (E9f368) and 45% (E5f157) (Figure 80). In the same way, Kaplan-Meier survival curves were constructed for the 6 AFLP markers for which the band-present phenotype corresponded to a susceptible allele. In those cases, the band-present phenotype was characterised by a decrease of the survival of between 24% (B12f243) and 40% (A1f150) (Figure 81).





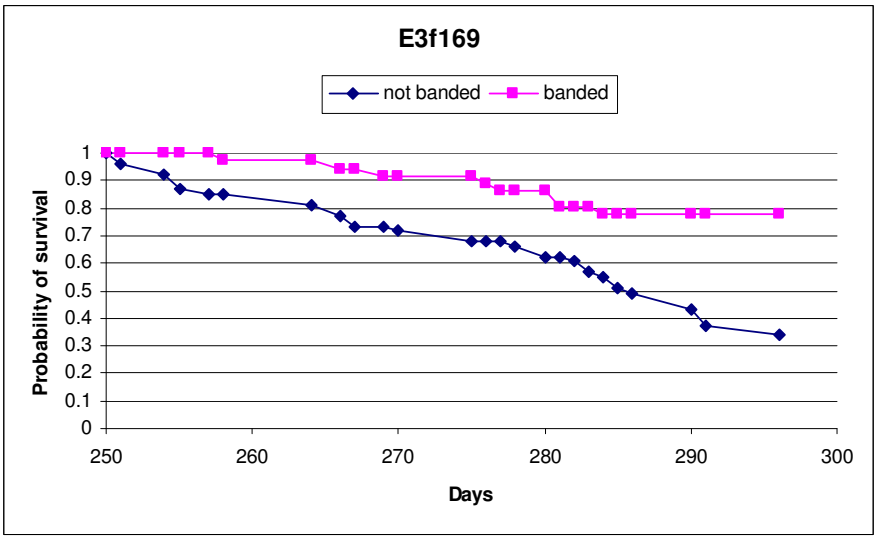
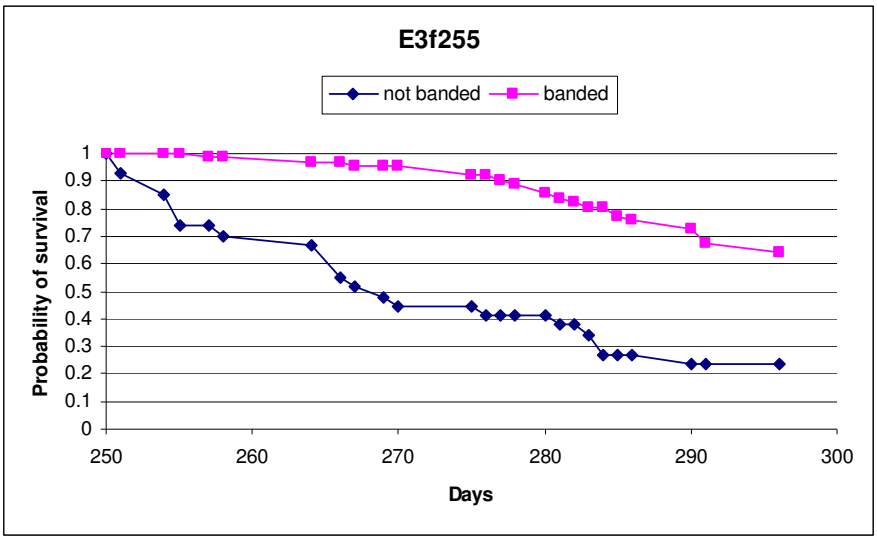
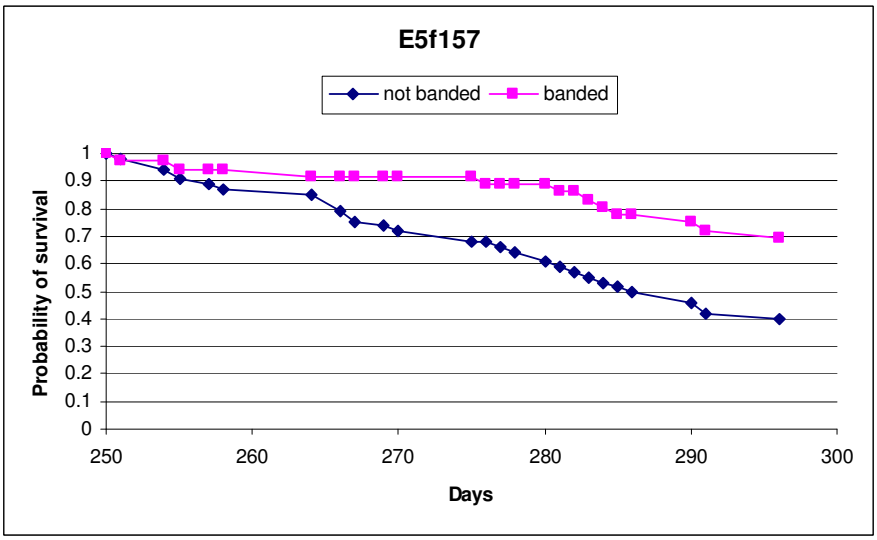


Figure 80-continued





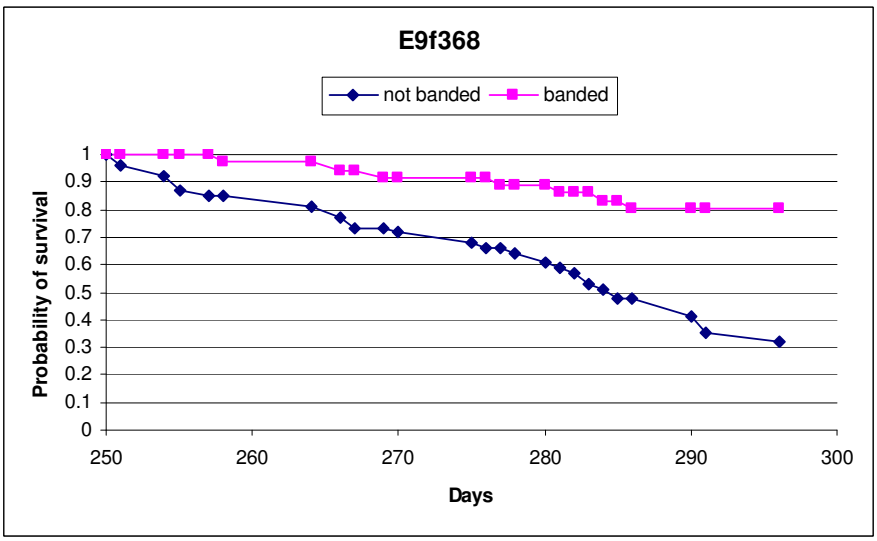


Figure 80-continued

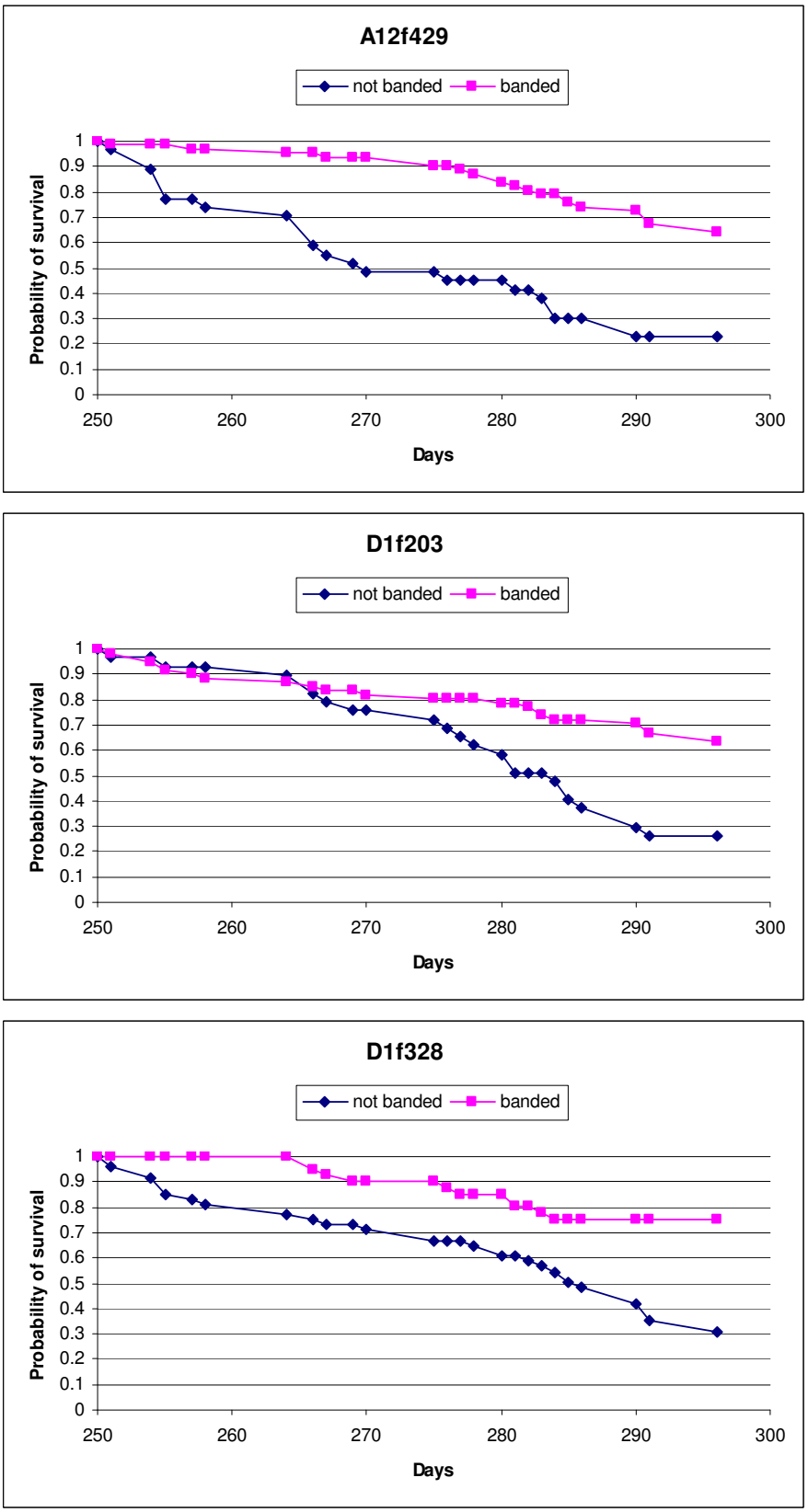
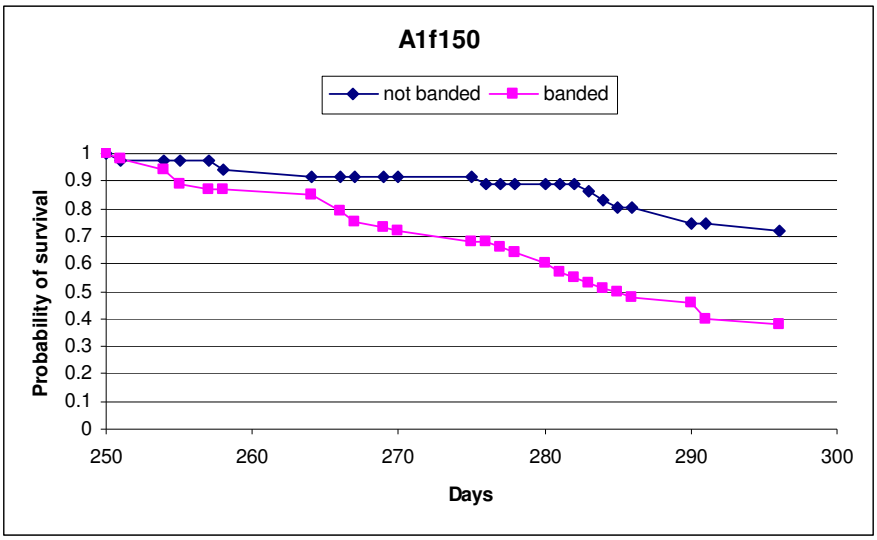
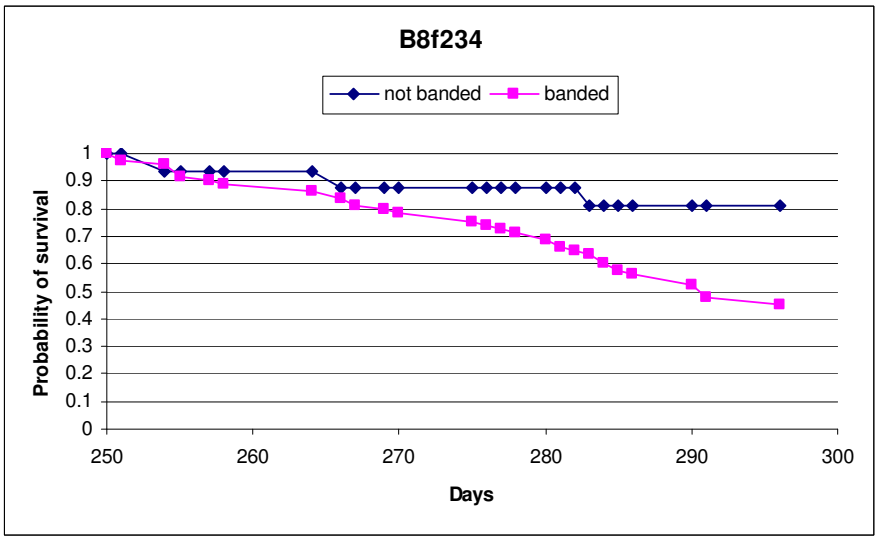


Figure 80. Kaplan-Meier survival curves for 9 AFLP markers for which the band-present phenotype corresponded to a resistant allele. Days: number of days after the beginning of the challenge experiment.





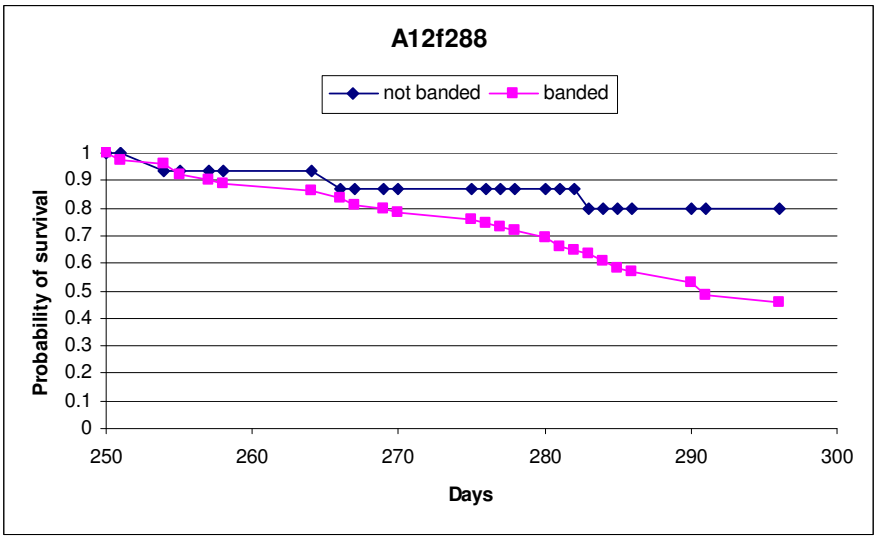


Figure 81-continued

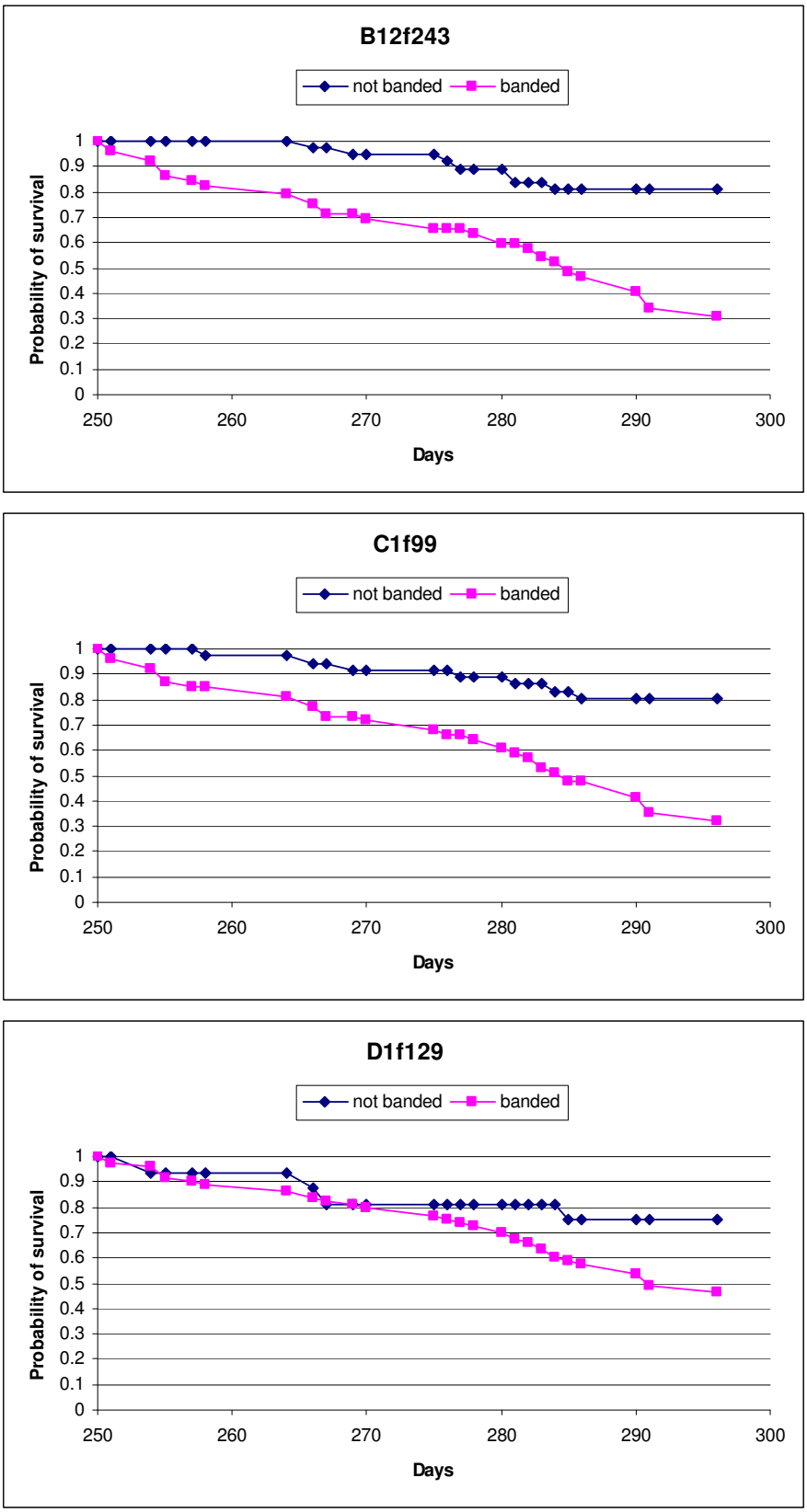


Figure 81. Kaplan-Meier survival curves for 6 AFLP markers for which the band-present phenotype corresponded to a susceptible allele. Days: number of days after the beginning of the challenge experiment.

IV.3. Genetic linkage mapping in the family OE.F2.05.04

The 410_7 parental genetic linkage map was based on the 17 microsatellites that were informative for that parent and 157 AFLPs segregating through this parent. The AFLPs consisted of 112 markers of type 1:1 (43 from 98AC703-29, 55 from W31 and 14 from both grand-parents) and 45 markers of type 3:1 (29 from 98AC703-29 and 16 from W31). The resulting map consisted of 127 markers (73.0%), comprising 16 microsatellites (on 17: 94.1%), 94 type 1:1 AFLPs (on 112: 83.9%), 17 type 3:1 AFLPs (on 45: 37.8%). Ten linkage groups were set up for the 410_7 map covering 465.6 cM (Figure 82). The sizes of the linkage groups ranged from 15.1 cM to 83.9 cM. The number of markers per linkage group varied from 5 to 25. The average distance between 2 loci ranged from 1.51 cM (G10_410_7) to 8.39 cM (G1_410_7), with an average spacing of 4.0 cM. The largest interval varied from 8.6 cM (G10_410_7) to 38.9 cM (G1_410_7) (Table 32). Distorted markers tended to cluster in specific linkage groups (G3_410_7, G4_410_7, G7_410_7, G8_410_7, G9_410_7 and G10_410_7). The observed map length was 465.6 cM for the 410_7 map. The estimated genome length was 553.37 cM according to method 4 of Chakravarti et al. (1991). The observed coverage was therefore 84.1 % for the 410_7 parental map.

Table 32. Length, number of markers, average spacing and largest interval between markers in linkage groups of the 410_7 parental map in *O. edulis* established with CriMap. Family OE.F2.05.04.

Linkage group	Length (cM)	No. of markers	Average spacing (cM)	Largest interval (cM)
G1_410_7	83.9	11	8.39	38.9
G2_410_7	70.2	10	7.8	33.3
G3_410_7	70.1	25	2.92	16.7
G4_410_7	59.6	23	2.71	17.0
G5_410_7	49.1	10	5.46	14.5
G6_410_7	45.2	12	4.11	19.4
G7_410_7	33.0	9	4.13	10.8
G8_410_7	19.9	11	1.99	12.2
G9_410_7	19.5	5	4.88	12.0
G10_410_7	15.1	11	1.51	8.6
Total	465.6	127		

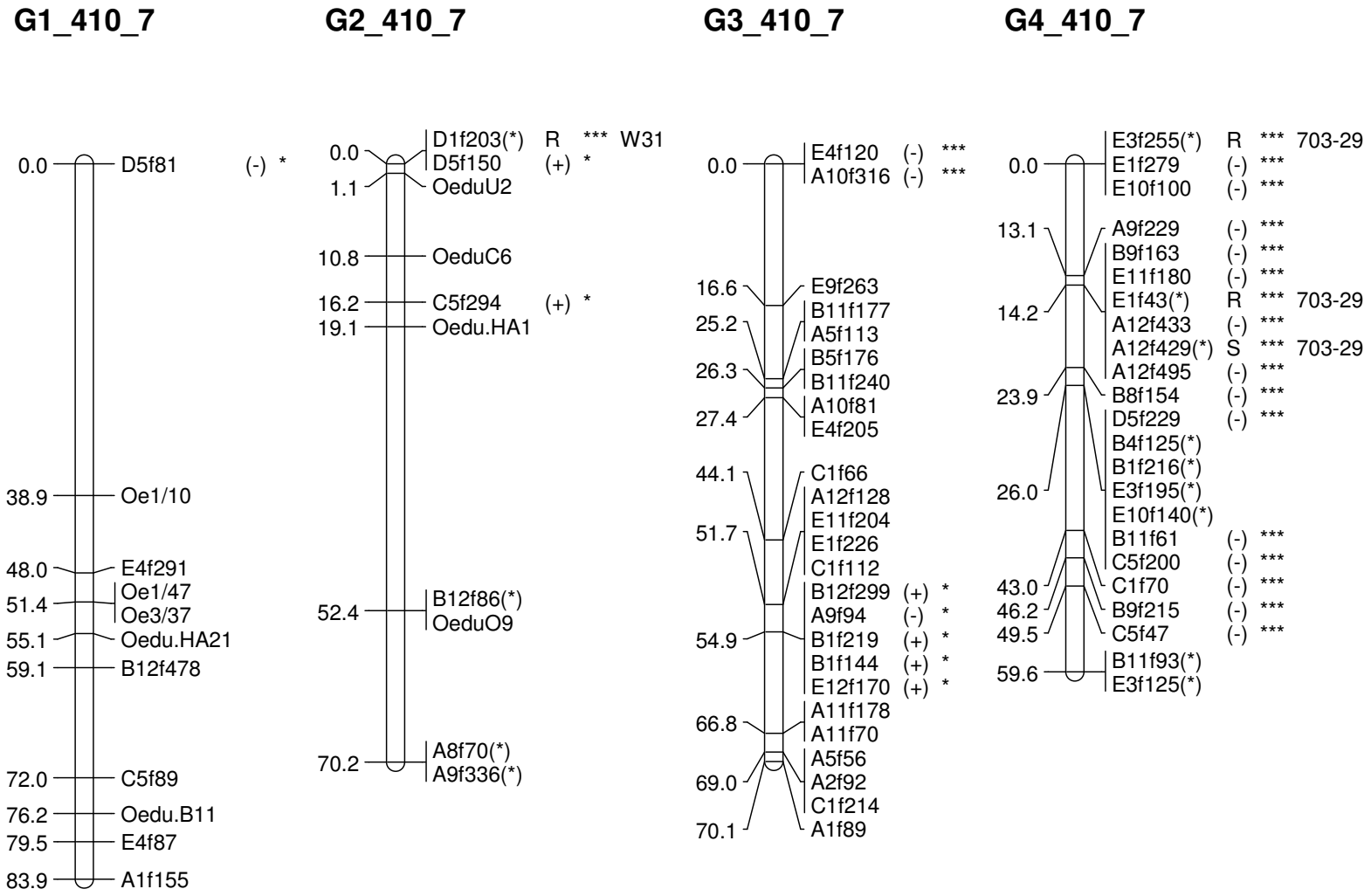


Figure 82-continued

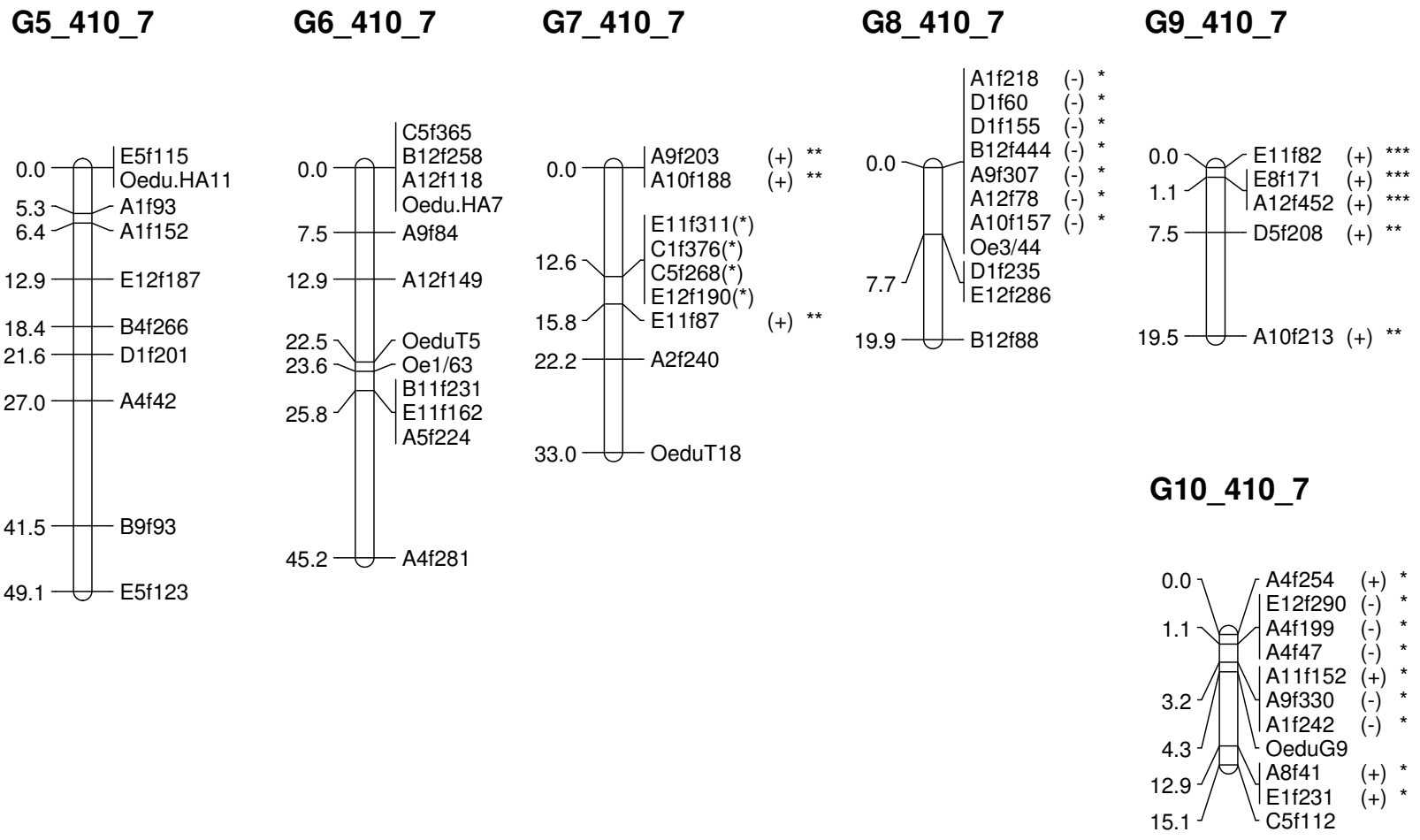


Figure 82. Microsatellite and AFLP-based linkage map of the flat oyster *O. edulis* in the mapping family OE.F2.05.04: 410_7 parental map obtained with CriMap, 127 markers, 466 cM. AFLP markers are labelled with the primer pair name followed by the letter "f" (for fragment) and a 3-digit fragment size in base pairs. Markers are indicated on the right; and absolute positions on the left (in Kosambi cM). On the right of AFLP locus name are specified the direction of the segregation distortion: towards a deficit (-) or excess (+) of null homozygotes with the statistical significance (* $p<0.05$, ** $p<0.01$ and *** $p<0.001$). R: resistant marker, S: susceptible marker (detected by the approach described in Moen et al., 2004a), with the statistical significance and grand-parental origin. Locus name immediately followed by an asterisk (*) correspond to 3:1 type AFLP.

The 410_8 parental genetic linkage map was based on the 18 microsatellites that were informative for that parent and 124 AFLPs segregating through this parent. The AFLPs consisted of 79 markers of type 1:1 (31 from 98AC703-29, 39 from W31 and 9 from both grand-parents) and 45 markers of type 3:1 (29 from 98AC703-29 and 16 from W31). The resulting map consisted of 98 markers (70.0%), comprising 14 microsatellites (on 18: 77.8%), 71 type 1:1 AFLPs (on 79: 89.9%), 13 type 3:1 AFLPs (on 45: 28.9%). Ten linkage groups were set up for the 410_8 map covering 386.7 cM (Figure 83). The sizes of the linkage groups ranged from 11.0 cM to 68.1 cM. The number of markers per linkage group varied from 4 to 19. The average distance between 2 loci ranged from 2.88 cM (G6_410_8) to 17.2 cM (G5_410_8), with an average spacing of 4.4 cM. The largest interval varied from 8.6 cM (G6_410_8) to 40.9 cM (G2_410_8) (Table 33). Distorted markers towards a deficit (-) or excess (+) in null homozygotes could be mapped on 6 different linkage groups: G1_410_8 [2 (+) and 1 (-)], G3_410_8 [4 (-) and 1 (+)], G4_410_8 [1 (-)], G6_410_8 [3 (+) and 3 (-)], G8_410_8 [2 (+)] and G10_8 [2 (+)]. Generally, distorted markers in the same direction and with the same level of statistical significance tended to cluster and to be mapped over a very short distance (e.g. 2 or 3 (+) markers in G1_410_8, G6_410_8, G8_410_8 or G10_410_8; 4 (-) markers in G3_410_8). For the 410_8 parental map, the observed map was 386.7 cM and the estimated genome length 556.91 cM. The observed coverage was therefore 69.4 % for the 410_8 map.

Table 33. Length, number of markers, average spacing and largest interval between markers in linkage groups of the 410_8 parental map in *O. edulis* established with CriMap. Family OE.F2.05.04.

Linkage group	Length (cM)	No. of markers	Average spacing (cM)	Largest interval (cM)
G1_410_8	68.1	19	3.78	13.1
G2_410_8	66.1	10	7.34	40.9
G3_410_8	61.5	13	5.12	20.4
G4_410_8	54.1	15	3.86	14.7
G5_410_8	51.7	4	17.2	35.0
G6_410_8	46.1	17	2.88	8.6
G7_410_8	45.0	6	9.0	24.0
G8_410_8	24.0	6	4.8	22.9
G9_410_8	13.2	4	4.4	9.9
G10_410_8	11.0	4	3.67	11.0
Total	98	386.7		

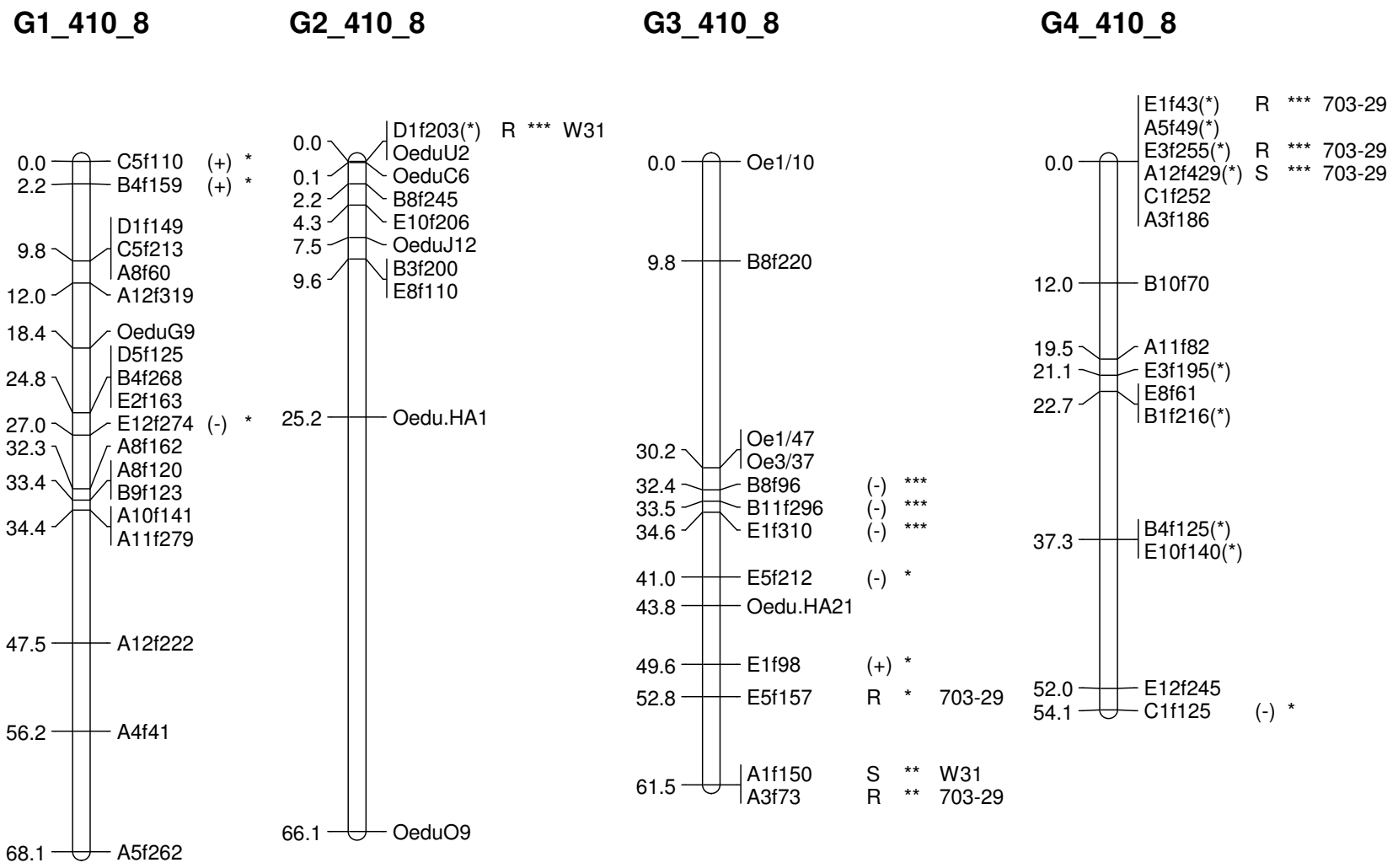


Figure 83-continued

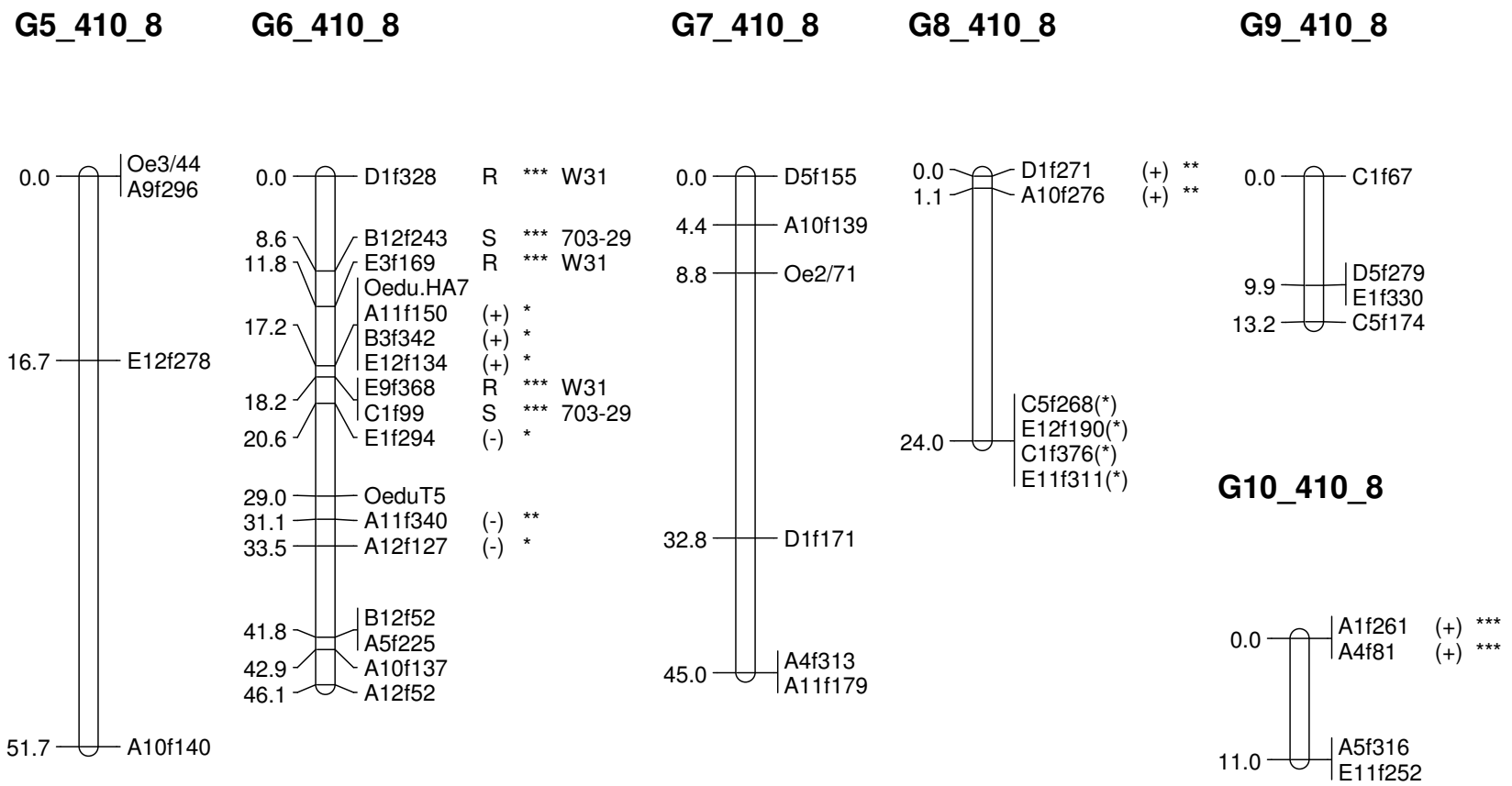


Figure 83. Microsatellite and AFLP-based linkage map of the flat oyster *O. edulis* in the mapping family OE.F2.05.04: 410_8 parental map obtained with CriMap, 98 markers, 387 cM. AFLP markers are labelled with the primer pair name followed by the letter "f" (for fragment) and a 3-digit fragment size in base pairs. Markers are indicated on the right; and absolute positions on the left (in Kosambi cM). On the right of AFLP locus name are specified the direction of the segregation distortion: towards a deficit (-) or excess (+) of null homozygotes with the statistical significance (* $p<0.05$, ** $p<0.01$ and *** $p<0.001$). R: resistant marker, S: susceptible marker (detected by the approach described in Moen et al., 2004a), with the statistical significance and grand-parental origin. Locus name immediately followed by an asterisk (*) correspond to 3:1 type AFLP.

Of the 7 markers that remained significant after the survival analysis and that segregated through the parent 410_7, 4 could be mapped on the 410_7 map, 3 resistant markers (band-present marker phenotype associated with a better survival) and 1 susceptible marker (band-present phenotype associated with a higher mortality):

- marker D1f203 on G2_410_7 (resistant, grand-parental origin: W31)
- marker E3f255 and E1f43 on G4_410_7 (resistant, grand-parental origin: 98AC703-29)
- marker A12f429 on G4_410_7 (susceptible, grand-parental origin: 98AC703-29)

The 3 markers on G4_410_7 were mapped in a 14 cM area and interestingly they have the same grand-parental origin, while the 4th marker (D1f203) mapped in another linkage group came from the second grand-parent.

For the 410_8 parent, of the 15 markers significant after the survival analysis, 12 of those could be mapped on the 410_8 map, 8 resistant markers and 4 susceptible markers. These 12 markers were distributed on 4 linkage groups:

- marker D1f203 on G2_410_8 (resistant, grand-parental origin: W31)
- E5f157 and A3f73 on G3_410_8 (resistant, grand-parental origin: 98AC703-29)
- A1f150 on G3_410_8 (susceptible, grand-parental origin: W31)
- E1f43 and E3f255 on G4_410_8 (resistant, grand-parental origin: 98AC703-29)
- A12f429 on G4_410_8 (susceptible, grand-parental origin: 98AC703-29)
- D1f328, E3f169 and E9f368 on G6_410_8 (resistant, grand-parental origin: W31)
- B12f243 and C1f99 on G6_410_8 (susceptible, grand-parental origin: 98AC703-29)

The 3 markers on G3_410_8 mapped in a 9 cM area; the markers on G4_410_8 clustered altogether and the 5 markers on G6_410_8 mapped in a 18 cM area. Moreover, it is interesting to note that resistant markers on the same linkage group came from the same grand-parent while the susceptible markers came from the other grand-parent (G3_410_8 and G6_410_8).

Thanks to the mapping of microsatellites in the two parental maps, as well as a few 3:1 type AFLPs, seven probable homology groups were found. When several common markers were mapped in the two parental maps, marker order could be compared. Multiple and parallel linkages were found for 4 of these homology groups (G3_410_8/G1_410_7, G6_410_8/G6_410_7, G2_410_8/G2_410_7 and G8_410_8/G7_410_7) (Figure 84).

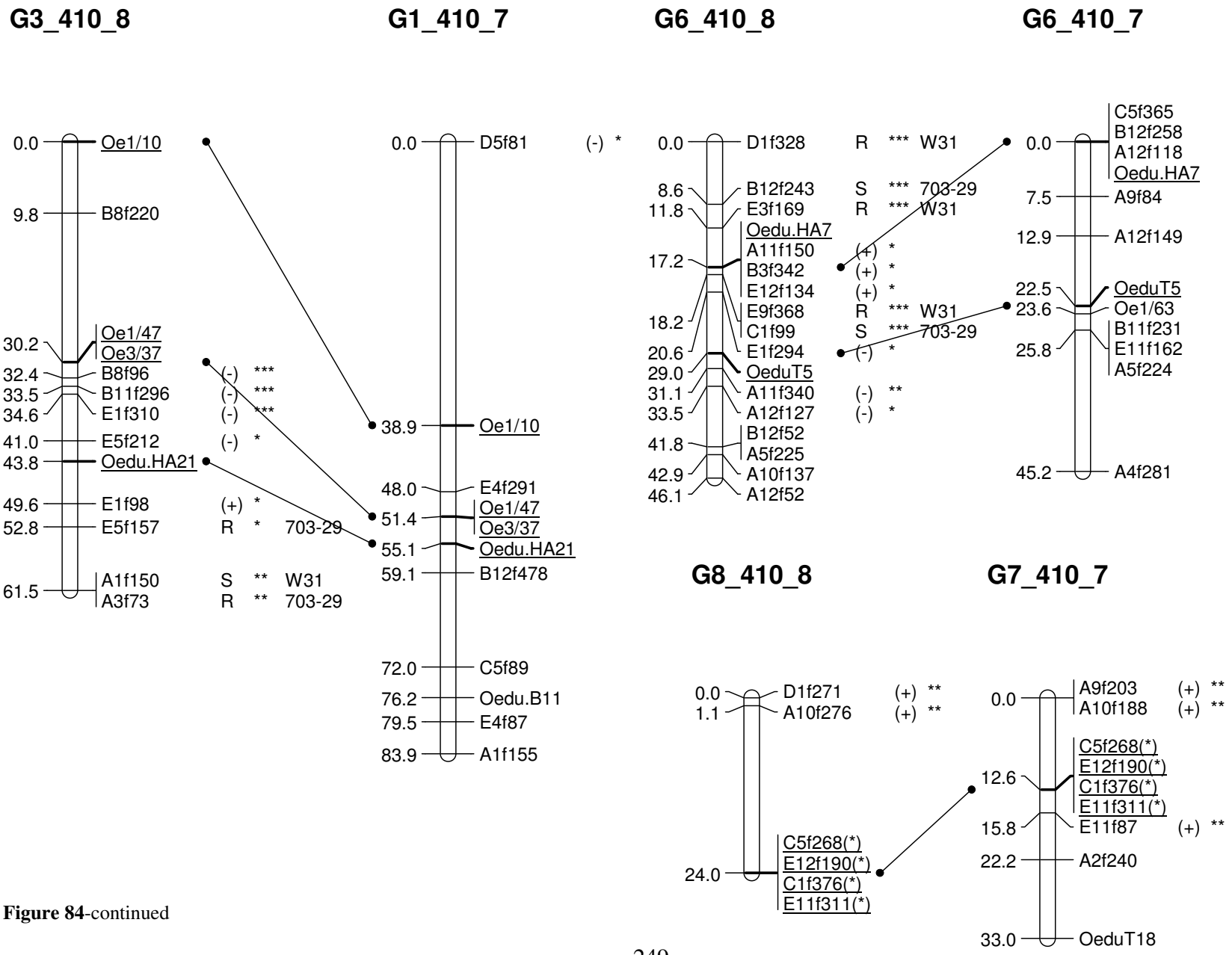


Figure 84-continued

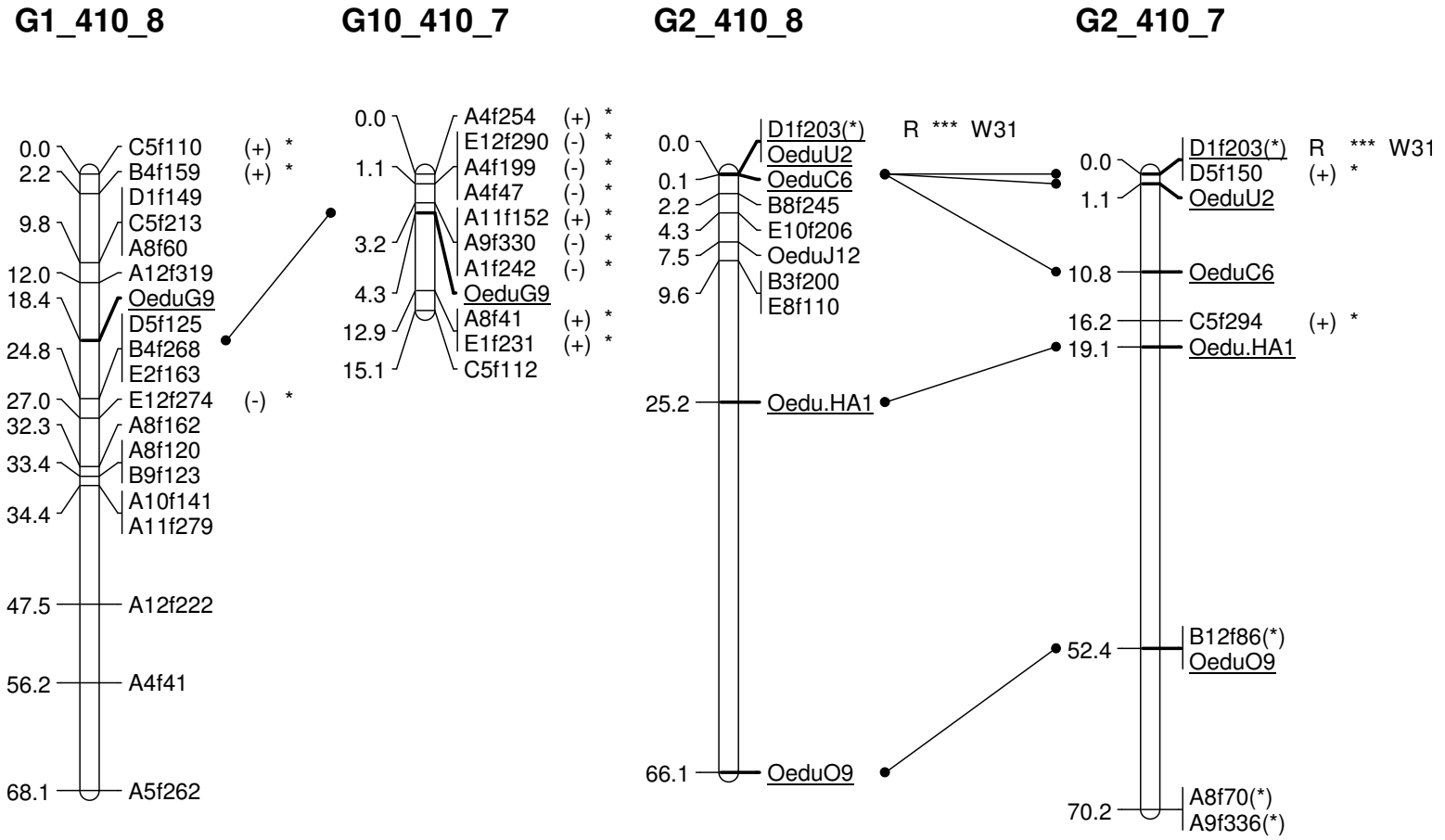


Figure 84-continued

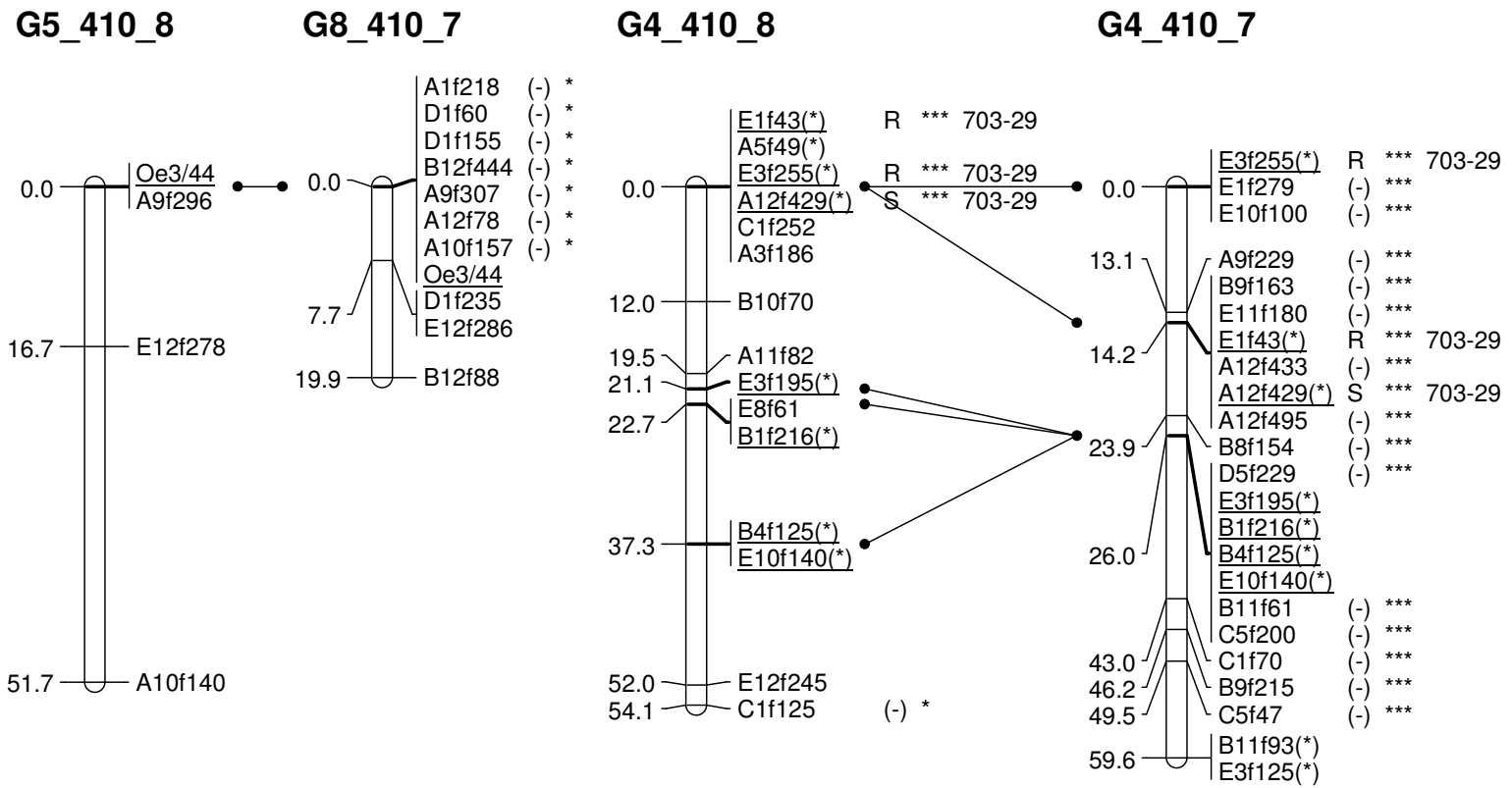


Figure 84. Seven pairs of homology groups found between the two parental maps 410_7 and 410_8 of the family OE.F2.05.04. See Figure 82 and 83 for the main abbreviations. Underlined markers are the common markers mapped in the two parental maps, on which the finding of homology groups was based. Lines were drawn between common markers to identify parallel linkages.

Two homology group pairs were identified based on the mapping of only one microsatellite in the two parental maps (G1_410_8/G10_410_7 and G5_410_8/G8_410_7). These homologies should be confirmed by the mapping of more common markers (Figure 84).

It was interesting to compare the mapping of the markers that were significant after the survival analysis in the two parental maps. The marker D1f203 mapped in the two parental maps at the same end of the linkage groups (G2_410_8 and G2_410_7), very close to microsatellite *Oedu*U2. Moreover, the 3 markers E1f43, E3f255 and A12f429 were mapped in the two parental maps, in the terminal part of the linkage groups: they were clustered at the end of group G4_410_8, while in G4_410_7 they were mapped in a 14 cM terminal area with two of them being clustered (E1f43 and A12f429) (Figure 84).

IV.4. QTL mapping in the family OE.F2.05.04

For the 410_7 parent, 4 linkage groups exhibited a significant paternal estimate, meaning that the two different gametes carried alleles of different effect for the QTL. The best estimate of location for the QTL was 0 cM for G2-410_7, 0 cM for G3_410_7, 24 cM for G4_410_7 and 8 cM for G6_410_7 (Figure 85). Only linkage group G2_410_7 attained significance at the chromosome-wide 0.01 level. The two linkage groups G4_410_7 and G6_410_7 attained significance at the chromosome-wide 0.05 level, but only just (Table 34).

Table 34. Results from fitting a single QTL for the parent 410_7. Threshold p 0.05 and threshold p 0.01 correspond to chromosome-wide significance thresholds at α=5% and 1% after performing 1000 permutations.

Linkage group	F ratios			Location (cM)	Paternal estimate (standard error)
	Threshold p 0.05	Threshold p 0.01	Observed		
G2_410_7	6.78	9.87	83.65	0	0.3574 (0.039)
G3_410_7	7.97	12.14	5.57	0	-0.134 (0.057)
G4_410_7	6.63	11.72	6.73	24	-0.1601 (0.062)
G6_410_7	6.65	10.65	6.65	8	-0.1329 (0.051)

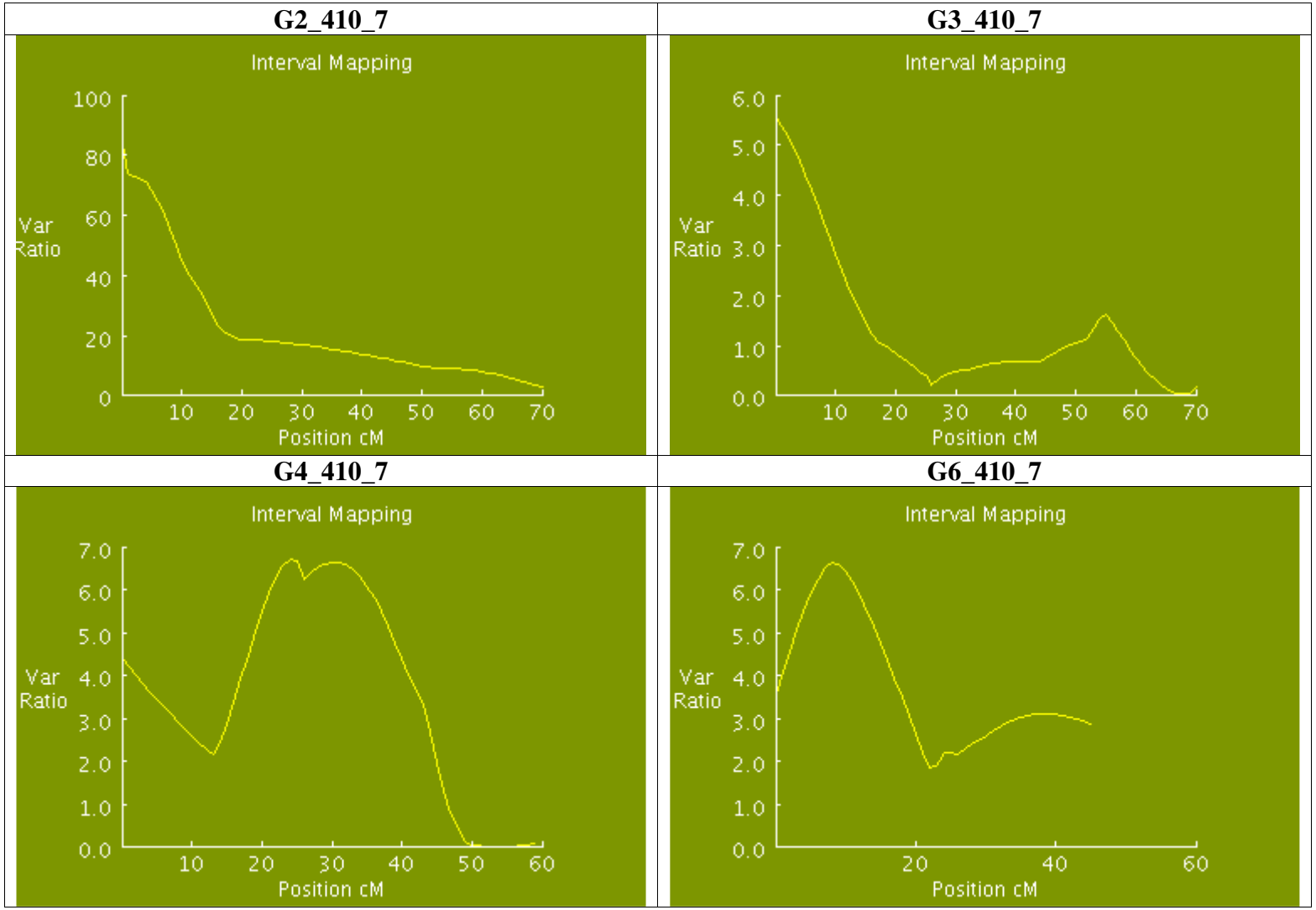


Figure 85. F ratios in 4 linkage groups of the 410_7 parental map, QTL being located at the position associated with the highest F ratio.

For the 410_8 parent, 2 linkage groups exhibited a significant maternal estimate, meaning that the two different gametes carried alleles of different effect for the QTL. The best estimate of location for the QTL was 61 cM for G3_410_8 and 17 cM for G6_410_8 (Figure 86). Linkage group G3_410_8 attained significance at the chromosome-wide 0.05 level, and G6_410_8 at the chromosome-wide 0.01 level (Table 35).

Table 35. Results from fitting a single QTL for the parent 410_8. Threshold p 0.05 and threshold p 0.01 correspond to chromosome-wide significance thresholds at α=5% and 1% after performing 1000 permutations.

Linkage group	F ratios			Location (cM)	Maternal estimate (standard error)
	Threshold p 0.05	Threshold p 0.01	Observed		
G3_410_8	7.22	12.08	8.17	61	-0.149 (0.052)
G6_410_8	7.96	12.14	22.19	17	0.229 (0.049)

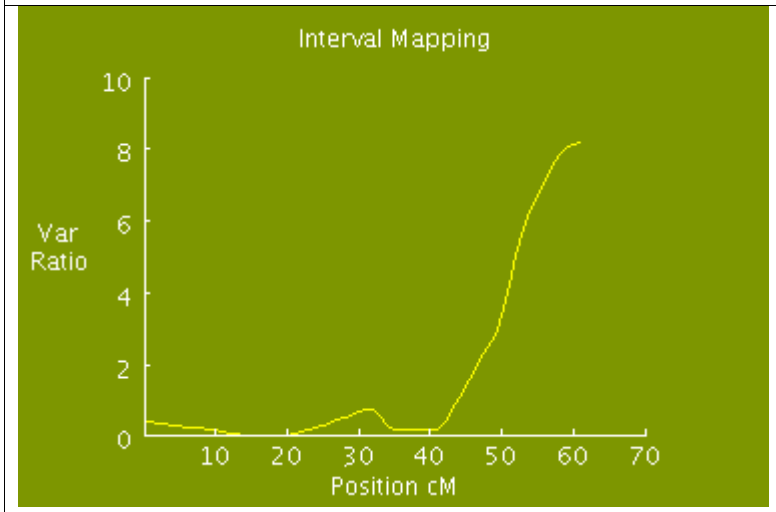
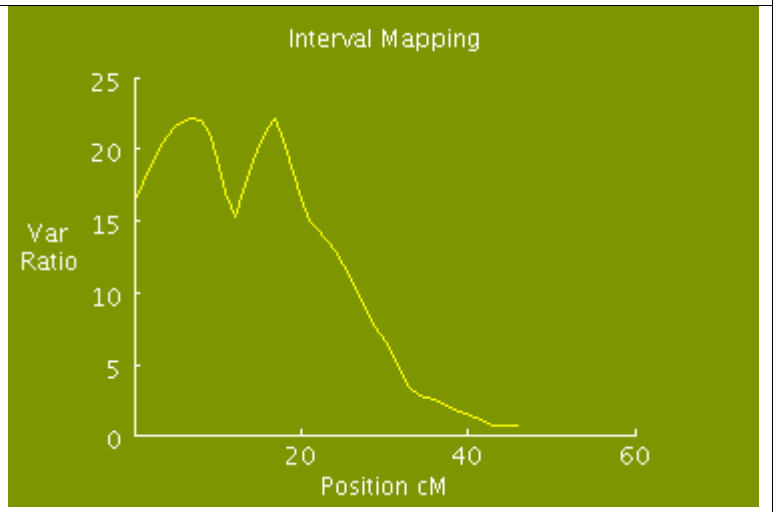
G3_410_8	G6_410_8
	

Figure 86. F ratios in 2 linkage groups of the 410_8 parental map, QTL being located at the position associated with the highest F ratio.

V- DISCUSSION

V.1. Segregation distortion: implications of the high genetic load

Relatively high segregation distortion was reported in the mapping family OE.F2.05.04. Overall, 25% of the markers were distorted, 47.4% of the microsatellites (9 out of 19) and 23.6% of the AFLPs (73 out of 309). The range of segregation distortion reported in this study was similar to the one reported in *C. gigas*, 20.9%-31% depending on the study (McGoldrick and Hedgecock, 1997; Launey and Hedgecock, 2001; Li and Guo, 2004).

The mapping family under study (OE.F2.05.04) came from crossing into a selected oyster strain that had been through a strong population bottleneck with a small effective number of breeders (Launey et al., 2001). The link between the relatively high level of segregation distortion and the inbreeding depression undergone during the selective breeding process suggests that the relatively high distortion in segregation patterns observed was due to a high genetic load in *O. edulis*, as previously stated by Bierne et al. (1998) in this species.

Moreover, the distribution of distorted markers was not random in the genetic linkage maps produced and tended to form clusters of distorted markers that were restricted to a few linkage groups (e.g. G3_410_7, G8_410_7 or G3_410_8; Figures 82 and 83). These clusters of distorted markers could therefore correspond to the location of potential deleterious genes in *O. edulis*, similarly to the mapping of potential deleterious genes in the rainbow trout (Young et al., 1998) or the Pacific oyster *C. gigas* (Li and Guo, 2004).

V.2. Linkage map and genome coverage

The genome coverage achieved was good, above 82% the 410_7 parental map, but less (69.4%) for the 410_8 parental map. Those genome coverage estimates compared favourably with the ones established in cupped oysters' species which were in the range of 70-90% depending on the study (Yu and Guo, 2003; Hubert and Hedgecock, 2004; Li and Guo, 2004).

Moreover, the number of linkage groups in the two parental maps 410_7 and 410_8 matched the haploid number of 10 chromosomes in this species (Thiriot-Quiévreux and Ayraud, 1982; Thiriot-Quiévreux, 1984b) (Figures 82 and 83).

However, more markers should be added to increase the genome coverage. Indeed, only 7 probable homology groups could be found in the mapping family, instead of 10 (Figure 84). Some of the linkage groups consisted of only two markers or spanned a small genetic distance (< 20 cM). Therefore, these groups may in fact belong to the same chromosome and may coalesce by adding more markers. These observations (less than 10 homology groups and small linkage groups) confirm the preliminary nature of the maps obtained.

V.3. *Bonamia* challenge experiment

A cohabitation experiment was chosen to transmit the disease from artificially infected wild oysters to a tested family, because it probably mimics more the natural mode of transmission of the disease. Indeed, challenge by injection would create stress and tissue lesion, and would bypass the natural pathway of entry of the parasite into the oyster.

The *Bonamia* challenge experiment successfully demonstrated that the disease was transmitted from the wild oysters to the tested oysters, with the first mortalities in the tested oysters occurring after 4 months of cohabitation (Figures 77 and 78). This compares favourably with the classical period of 3 to 5 months for the transmission period (Tigé and Grizel, 1984; Grizel, 1985). However, mortalities occurred at a higher rate in the family OE.F2.05.04 (oysters of 8 month-old at the beginning of the challenge experiment) than in the family OE.F2.04.45 (oysters of 20 month-old at the beginning of the challenge experiment) (χ^2=24.34, p<0.001, 4 d.f.). Moreover, the level of infection was significantly lower in the family OE.F2.04.45 (χ^2=20.87, p<0.001, 1 d.f.). This result was inconsistent with the findings of Culloty and Mulcahy (1996) that 2 years is the critical age for the disease development. This strengthens that age may not be the key factor in development of the disease, but that the disease could be better related to factors other (size) than the age of the oyster (Caceres-Martinez et al., 1995). In our study, the mean size of the dead oysters of the family OE.F2.05.04 was 4.9 cm. A recent study followed the

prevalence and intensity of infection in young prespawning oysters (2-3 month old to 18 month-old) by heart smear and PCR techniques and showed that such young oysters were susceptible to infection (Lynch et al., 2005). These authors analysed the level of infection to the parasite but did not report any mortality due to *Bonamia*. Our study reports mortalities due to *B. ostreae* in one year old oysters and demonstrates susceptibility of young oysters to the disease in our artificial hatchery conditions.

The most probable discrepancy in the development of the disease between the 2 tested families (OE.F2.04.45 and OE.F2.05.04) is a different genetic background. Indeed, both families originated from the cross between a wild oyster and an oyster derived from the selective breeding programme to bonamiasis (oyster strains S85 and S89). However, only the family OE.F2.05.04 was directly derived from a selected family (family 98AC703 produced in 1998) (Figures 70 and 72) that had been tested for its resistance to bonamiasis both in the field and by injection. In contrast, the family OE.F2.04.45 originated from a 6^{th} generation inbred line that showed no mortality in the field but whose resistance was not tested by inoculations (Figures 42 and 43). Moreover, because the two families issued from a cross involving a wild oyster, chance probably played a role in this difference of reaction of the two families concerning the development of bonamiasis because the status of resistance/susceptibility to the disease of the wild oyster parents was unknown.

V.4. QTL mapping of resistance/susceptibility to a disease

The development of genetic linkage maps is particularly useful for the mapping of quantitative trait loci (QTLs) and for marker-assisted selection (MAS). Several studies have highlighted the potential for marker assisted selection in breeding programmes in fisheries and probably shellfisheries in the future (Ward et al., 2000; Perry et al., 2001; Liu and Cordes, 2004). MAS has a huge potential in aquaculture breeding programme, especially for traits difficult to phenotype and would reduce the time between two generations of selection. But so far no successfully applied MAS has been reported in either fish or shellfish species.

Although disease resistance generally seems to have a low heritability in some species (Gjedrem, 2000) it is nevertheless an ideal trait for the application of marker-assisted selection (MAS), due to the economic significance of high survival in aquaculture. Traditionally, the QTL mapping approach was designed for continuously variable quantitative traits. However, it has been shown that this analysis is robust for binary traits, such as resistance/susceptibility to a disease (death/alive trait coded as 0 or 1) (Visscher et al., 1996). Therefore, we decided to couple the single marker approach with a traditional QTL mapping approach in order to compare the results obtained with these two alternative approaches. The finding of QTLs was achieved by using a multistage testing strategy (developed by Moen et al., 2004a) and a regression interval mapping (QTL express software, Seaton et al., 2002).

The multistage testing strategy was a very powerful and robust way to identify QTLs of resistance/susceptibility to a disease and allowed us to identify 15 probable AFLP markers linked to genes of resistance (for 9 of them) or susceptibility (for 6 of them) to the disease. Because this test involved three successive tests, a transmission disequilibrium test in the affected oysters, a Mendelian segregation test in the whole dataset (affected and non-affected oysters) and a survival analysis, we believe that this approach is more robust than the sole survival analysis and reduces the detection of false positives (markers considered as linked to disease resistance or susceptibility when in fact they are not). Its main advantage is its simplicity and that it does not require the use of a genetic map. It is therefore very interesting for obtaining preliminary results.

The QTL mapping approach chosen was a regression interval mapping (interval mapping based on least-squares regression methods) (Haley and Knott, 1992; Martinez and Curnow, 1992). Indeed, an interval mapping approach is based on information from 2 linked flanking markers and has been shown to be more powerful than a single marker analysis particularly for medium-density maps (with markers around 20-35 cM apart) and to increase the accuracy of parameters estimation (Darvasi et al., 1993; Tanksley, 1993). The analysis was based on a single full-sibs family experimental design (and not a F2 analysis) because the mapping family was a three-generation outbred family whose grand-parents were not issued from different genetic lines fixed for different alleles at the QTL. The analysis was interpreted in terms of paternal and maternal components, i.e. whether the 2 gametes of each parent carried alleles of different effect

for the QTL. However, no estimation of the interaction component (and therefore dominance of the QTL) could be performed because of the lack of anchor loci (e.g. microsatellites). Furthermore, a consensus map could not be built because most markers were segregating in only one of the two parents and the estimation of the interaction component implies that the map is the same in the two parents. Moreover, because of the lack of power of the experiment (92 progeny genotyped), a one-QTL model was fitted but not a two-QTL model. A two-QTL model would try to fit two QTLs on the linkage group, and would give the locations and estimates of the two QTLs. The comparison between the two models (one-QTL versus two-QTL) will determine which model is more appropriate. The results obtained after fitting the one-QTL model should be interpreted with caution because the role of neighbouring QTLs in biasing the estimation of location and gene effect of a QTL has been widely assessed (Lander and Botstein, 1989; Haley and Knott, 1992; Martinez and Curnow, 1992; Jansen, 1993): the results of the one-QTL model can be misleading when there are in fact 2 linked QTLs segregating on the same linkage group. Therefore, our experimental design was not optimal for QTL mapping and there is a need to add more codominant markers in future to optimise experimental design.

Despite the limitations of the study, interesting and promising results were obtained and it is clear that this study has made a big step towards the finding of resistance/susceptibility to bonamiasis. Indeed, several potential markers of interest were found and more importantly there was a good concordance between the results obtained after the multistage testing strategy, the genetic mapping approach and the QTL mapping approach itself. Indeed, significant markers after the survival analysis tended to cluster or were restricted to a few groups: in the 410_7 map, 3 markers were mapped in a 15 cM area in G4_410_7 group and a fourth marker was mapped in G2_410_7 group (Figure 82); in the 410_8 map, 3 markers were mapped in a 9 cM area in G3_410_8, 5 markers were mapped in a 17 cM area in G6_410_8, 3 markers were clustered at the end of G4_410_8 and one marker was mapped in G2_410_8 (Figure 83). Moreover, the significant QTLs found after the regression interval mapping approach were mapped in the same area as the markers that were significant after the survival analysis. Even if these results need to be confirmed and strengthened, the concordances were promising and suggest good confidence can be placed on the potential QTLs identified.

Several studies have reported the location of QTLs for disease resistance in rainbow trout, based on the classical approach for QTL mapping using interval mapping, the ANOVA-based approach, or Bulk Segregant Analysis (BSA) (Palti et al., 1999; Ozaki et al., 2001; Rodriguez et al., 2004). Disease resistance is of particular interest for the flat oyster that has suffered such a huge decline from parasitic diseases and we can foresee that MAS for disease resistance could be an important tool in the regeneration of oyster aquaculture. The results obtained in this study, even if preliminary, are promising and represent a first step towards MAS in the flat oyster. However, before implementation of MAS in a selective breeding programme, the role of epistasis and genomic background should be assessed (Danzmann et al., 1999; Ozaki et al., 2001; Perry et al., 2001; Perry et al., 2003).

V.5. Future work

The addition of codominant markers (such as microsatellites or SNPs) is critical to increase the accuracy of the genetic maps obtained and to allow the establishment of a consensus map. Moreover, in terms of QTL mapping, additional markers will also increase the power of detection of the QTLs and increase the accuracy of the estimation of the QTL effects. The dataset should be increased by genotyping more oyster progeny from the same mapping family, which should help resolving ambiguous map positions of markers that are closely linked and will allow the finding of more QTLs (of small to medium effect).

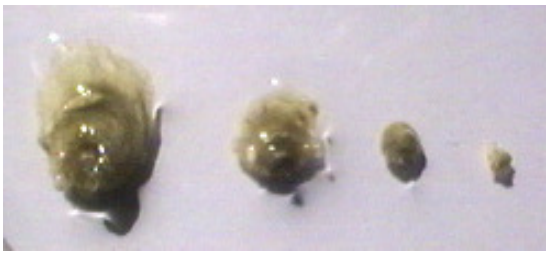
Another challenging experiment would be to confirm the finding of QTLs in other families, because the genetic background is a very important factor. The mapping of more codominant markers will therefore help in the location of the QTLs and comparison of QTLs found in different families.

It could be interesting to perform trials to assess the usefulness of selecting oysters based on their genotypes at the markers linked to the QTLs, i.e. to assess the potential for MAS. This could be tested by performing a first generation of selection and compared the gains with or without the help of information gained from QTL markers.

Finally, before implementation of MAS in *O. edulis*, fine QTL mapping should be achieved in order to restrict the region of interest to a more narrow area. A further step would be to go from the QTL down to candidate genes. The mapping of potential genes involved in the resistance to the disease that were found after performing a SSH (Suppression Subtractive Hybridisation) library would be of prime interest. Such candidate genes are currently under study by Benjamin Morga (PhD student at Ifremer, La Tremblade, France). Their inclusion into the genetic map would help to corroborate QTLs with candidate genes and would represent a further step into the understanding of the genetic component of the resistance/tolerance of *O. edulis* to *B. ostreae*.

CHAPTER 6

Inbreeding depression and genetic load in *Ostrea edulis*



I- INTRODUCTION

Self-fertilisation that occurred in the mapping family OE.F2.04.45 was unexpected because to our knowledge it is the first time that this phenomenon has been demonstrated in the flat oyster *Ostrea edulis*. This complicated our project because this family was initially chosen, based on DNA polymorphism, as the mapping family and was part of the *Bonamia*-challenge experiment. Indeed, our initial aim was to establish a first genetic linkage map in *O. edulis* and then to look for QTLs of resistance to bonamiasis by using a single mapping family. Unfortunately, the mapping family chosen was derived from a self-fertilisation of an F1 parent, was proved unsuitable for genetic mapping and was therefore discarded for QTL analysis due to the lack of informativeness.

Inbreeding corresponds to the mating between closely related individuals. Self-fertilisation is an extreme case of inbreeding. In a population that is inbred, the frequency of homozygotes is increased, whereas the frequency of heterozygotes is reduced relative to the Hardy-Weinberg expectations. The coefficient of inbreeding, *f*, is the probability that the two homologous alleles in an individual are identical by descent (IBD), i.e. these two alleles are copies of one particular allele possessed by a common ancestor. Inbreeding depression relates to the decline of fitness due to inbreeding and can be related to the expression of deleterious genes that were hidden in the heterozygous state. Indeed, the overall increase of homozygosity in the genome due to inbreeding also results in an increase of the homozygosity of recessive deleterious genes.

Genetic load corresponds to the reduction in fitness due to the presence of recessive deleterious alleles maintained by a selection-mutation balance and the segregation of homozygotes when there is a heterozygote advantage (Hedrick, 2000). A high genetic load has previously been reported in bivalves (McGoldrick and Hedgecock, 1997; Bierne et al., 1998; Launey and Hedgecock, 2001). This huge potential for genetic load in bivalve species is in accordance with the Elm-oyster model (Williams, 1975): high fecundity associated with high mortality in the early stages favours individual variation in fitness and in traits correlated with fitness, leading to correlation between fitness-related traits and heterozygosity at molecular loci

(Alvarez et al., 1989; David et al., 1995; Bierne et al., 1998; Launey and Hedgecock, 2001). Inbreeding depression and genetic load are therefore interrelated concepts.

Considerable effort was put into the self-fertilised family: genotyping of 20 microsatellites and 60 AFLP primer pairs for 48 F2 progeny, follow-up of growth over a 5 month period (prior the challenge experiment) and follow-up of mortality and parasite level of infection during the 6-month *Bonamia*-challenge experiment. Therefore, we decided to use the self-fertilised family (OE.F2.04.45) to investigate genetic load and inbreeding depression in *O. edulis*. This part of the study aimed to:

- compare segregation distortion for two families having different coefficients of inbreeding. Two families were compared, OE.F2.04.45 derived form a self-fertilisation of an F1 parent (f=0.5) and OE.F2.04.63 derived from a fullsib crossing (f=0.25). Unfortunately, a family derived from two unrelated oyster parents was not available for comparison.

- investigate the loss of heterozygosity depending on the level of inbreeding and its implications in terms of the production of homozygous inbred lines. The same two families were compared in terms of observed and expected heterozygosity.

- investigate the relationship between growth and multi-locus heterozygosity (MLH). This was done in the self-fertilised family, by using the 5-month growth data and the microsatellite genotypes (20 loci) that were initially obtained in the context of the genetic and QTL mapping experiments. A positive correlation between MLH and life history traits such as growth or survival has already been reported in *O. edulis* based on allozymes (Alvarez et al., 1989; Launey, 1998) and four microsatellite markers (Bierne et al., 1998).

II- MATERIAL AND METHODS

II.1. Oyster families

Two families were considered:

- an F2-family issued from a biparental cross between 2 F1 parents: family OE.F2.04.63 (see Figures 42 and 43 for details)
- an F2-family issued from a self-fertilisation of an F1 parent: family OE.F2.04.45 (see Figures 42 and 43 for details)

II.2. Molecular markers

Molecular markers consisted of 20 microsatellites and 60 different AFLP primer combinations (see Chapter 2 for protocols and Table 12).

II.3. Comparison of segregation distortion

Segregation distortion analysis was achieved by a χ^2 goodness of fit statistical test as explained in part III.3. of chapter 4 for the microsatellites and the AFLPs. Segregation distortion was compared for the two families OE.F2.04.45 (n=48) and OE.F2.04.63 (n=48) by performing a G-test of independence. Then, for the distorted microsatellites only, segregation distortion was interpreted in terms of Identical By Descent (IBD) status of the F2 individuals, meaning that IBD homozygotes have two copies of the allele coming from the same grand-parent. For example, consider two grand-parents whose genotypes at a particular microsatellite are 156/160 for the first grand-parent and 152/164 for the second grand-parent; and two parents F1 whose genotypes are 156/164 and 156/160. Therefore, in the F2 progeny, 4 equiprobable genotypes could be observed: 156/156, 156/160, 156/164 and 160/164. The F2 progeny of genotype 156/156 will be considered as homozygotes IBD because the two copies of allele 156 come from the first grand-parent.

II.4. Comparison of heterozygosity

Observed and expected heterozygosity were computed for each of the 20 microsatellites for the two families OE.F2.04.45 (n=48) and OE.F2.04.63 (n=48). Expected heterozygosity could be inferred from the genotypes of the parents: for example, if the genotypes of the two parents at a particular locus are 156/156 and 156/160, then 2 equiprobable genotypes could be observed in the progeny (156/156 and 156/160), the expected heterozygosity would be 0.5. The overall observed and expected heterozygosity were compared by a χ^2 goodness of fit statistical test within each of the two families.

II.5. Growth/heterozygosity relationship in the family OE.F2.04.45

In order to assess an eventual growth/heterozygosity relationship, 550 oysters of the family OE.F2.04.45 were individually labelled and placed in 5 different raceways, 5 trays per raceway, at a density of 25 oysters per tray. Oysters numbered OE.F2.04.45-1 to 125 were disposed into the Raceway 2, OE.F2.04.45-126 to 250 into Raceway 1, OE.F2.04.45-251 to 375 into Raceway 3, OE.F2.04.45-376 to 500 into Raceway 4 and OE.F2.04.45-501 to 550 into Raceway 5. We tried to control experimental conditions to avoid raceway-due biases: water flow was set up at 150-200 $l.h^{-1}$ for each raceway in order to have the same food supply for each raceway (phytoplankton was directly mixed upstream with the sea water) and trays within raceway were changed position at each cleaning. The 550 oysters were individually measured (total length) and weighed (after air drying the external shell for 2 hours) every month for 5 months: the 26/07, 30/08, 03/10, 09/11 and 19/12/2005. At the end of the experiment, 48 oysters were randomly chosen from the 5 raceways for DNA extraction from gill tissue and genotyping at the 20 microsatellites.

III- RESULTS

III.1. Comparison of segregation distortion in OE.F2.04.45 and OE.F2.04.63

For OE.F2.04.45, 17% of microsatellites and 50% of AFLPs segregated according to Mendelian rules and for OE.F2.04.63, 41% of microsatellites and 69% of AFLPs (Figure 87).

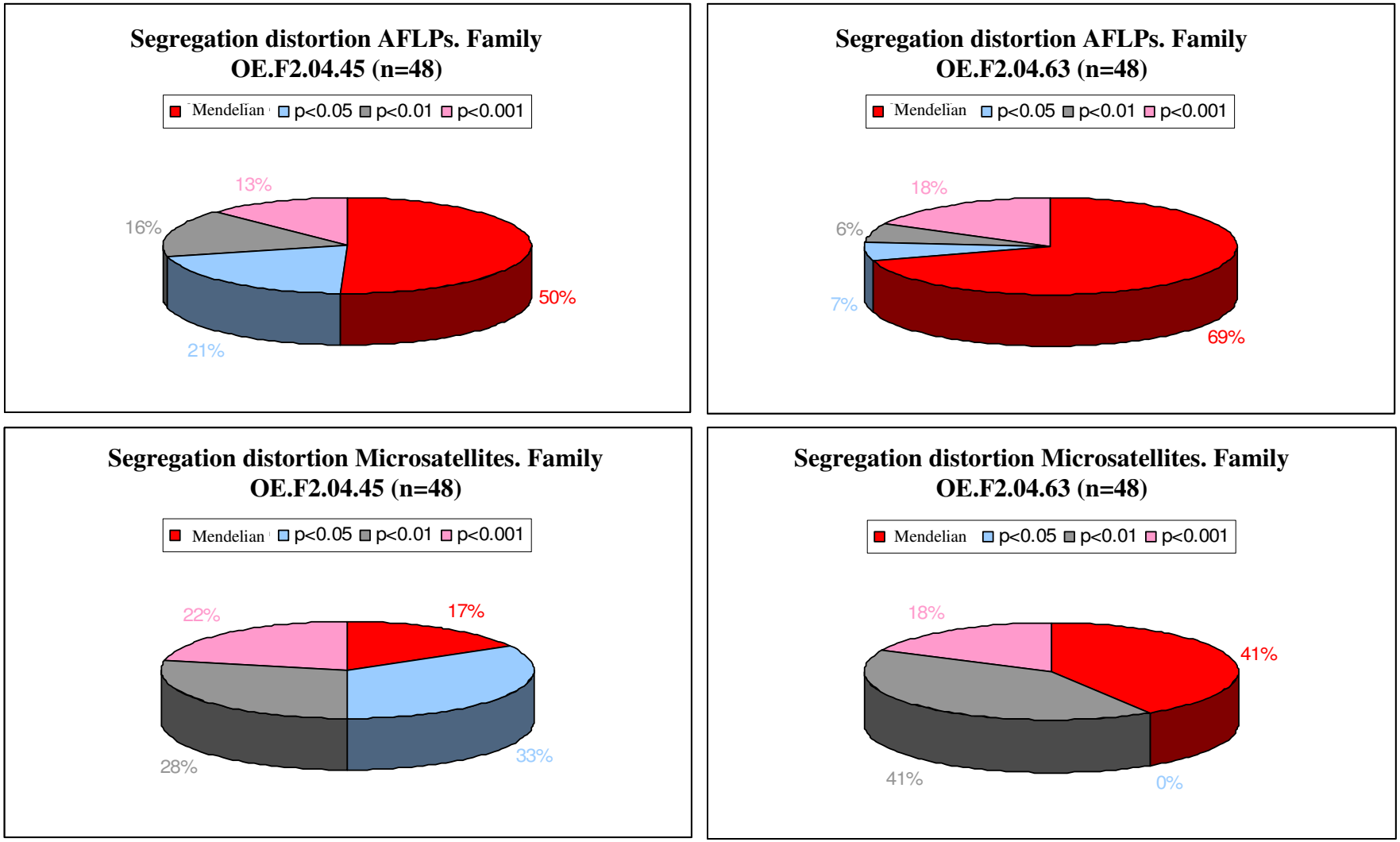


Figure 87. Comparison of segregation distortion for microsatellites and AFLPs in two families, OE.F2.04.45 and OE.F2.04.63. "Mendelian" correspond to markers that were segregating according to Mendelian rules; p<0.05, p<0.01 and p<0.001 to markers that showed distortion in their segregation ratio at a statistical significance of 5%, 1% and 1‰.

A G-test of independence revealed no significant difference between the two families in the segregation distortion at microsatellite markers (p=0.109). However, segregation distortion was significantly higher in the family OE.F2.04.45 for AFLP markers (p<0.001).

For OE.F2.04.45, all the significant departures from Mendelian segregation but marker Oe1/47 were due to a deficiency of one (or two) of the two IBD homozygous genotypes (AA or BB). The case is particular for Oe1/47 because the two grand-parents were 266/266 and 268/268: it was not possible to determine the IBD status at this locus (Figure 88). Moreover, χ^2 goodness of fit tests were achieved to check that segregation distortions were due to IBD homozygote deficiencies rather than heterozygote excess, by applying a goodness of fit test for the two genotypic classes that did not seem to depart from the expectation, test AB: BB (2: 1) or test AB: AA (2:1). Markers J12, U2, O9, G9, Oe2/71, HA1 and C6 exhibited a significant deficiency in one of the two homozygote IBD because the test AA: AB: BB was highly significant but the test AB: BB was not significant meaning that only one of the three genotypic classes (AA) departed significantly from its expectations. On the contrary, marker T5 was deficient in the homozygote IBD (BB). Markers HA7 and B11 were significantly deficient for the two homozygotes IBD (AA and BB). Markers Oe1/63 and HA21 were associated with both an heterozygote excess and a deficiency in the two homozygotes IBD (Table 36).

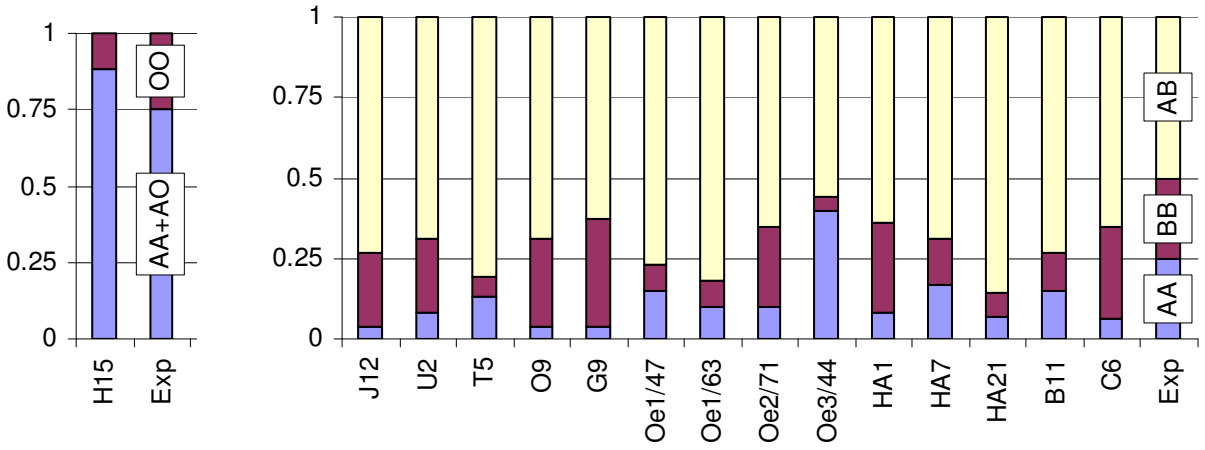


Figure 88. Relative proportions of microsatellite genotypes in 15 cases (out of 18) for which there was a significant departure (p<0.05) from expected Mendelian segregation ratios for the family OE.F2.04.45. Two types of segregation are shown, as indicated by the far-right bar of each type labelled "Exp" which shows the expected Mendelian proportions of each genotype. O denotes a null allele. AA corresponds to homozygotes IBD for the L002-55 grand-parent; BB to homozygotes IBD for the W120 grand-parent.

Table 36. χ^2 goodness of fit tests achieved in the family OE.F2.04.45 to test for Mendelian segregation (Test AA: AB: BB) and for deficiency in homozygote IBD (Test AB: BB or AB: AA) for microsatellites. Only microsatellites that were distorted after the Mendelian segregation test are shown in this table. NS: p>0.05; *: p<0.05; **: p<0.01; ***: p<0.001. Classes AA, AB and BB refer to Figure 88. Obs: observed number of F2 individuals in each class. IBD: identical by descent.

Marker	Test AA: AB: BB (1: 2: 1)			Test AB: BB (3: 1) or AB: AA (3: 1)			Conclusion
		Obs	Chi^2-value		Obs	Chi^2-value	
J12	244/244 (BB) 244/248 (AB) 248/248 (AA)	11 35 2	13.458 (***)	AB: BB	35: 11	0.03 (NS)	Deficiency in homozygote IBD (AA) (L002-55)
U2	170/170 (BB) 170/174 (AB) 174/174 (AA)	11 33 4	8.792 (**)	AB: BB	33: 11	0 (NS)	Deficiency in homozygote IBD (AA) (L002-55)
T5	152/152 (BB) 142/152 (AB) 142/142 (AA)	3 39 6	19.125 (***)	AB: BB AB: AA	39: 3 39: 6	7.143 (**) 3.267 (NS)	Deficiency in homozygote IBD (BB) (W120)
O9	146/146 (BB) 146/164 (AB) 164/164 (AA)	13 33 2	11.792 (***)	AB: BB	33: 13	0.261 (NS)	Deficiency in homozygote IBD (AA) (L002-55)
G9	214/214 (BB) 210/214 (AB) 210/210 (AA)	16 30 2	11.167 (***)	AB: BB	30: 16	2.348 (NS)	Deficiency in homozygote IBD (AA) (L002-55)
Oe1/47	268/268 (BB) 266/268 (AB) 266/266 (AA)	4 37 7	14.458 (***)	AB: BB AB: AA	37: 4 37:7	5.081 (*) 1.939 (NS)	Deficiency in homozygote IBD (BB)
Oe1/63	102/102 (BB) 102/106 (AB) 106/106 (AA)	4 39 5	18.792 (***)	AB: BB AB: AA	39: 4 39: 5	5.651 (*) 4.364 (*)	Significant excess of heterozygotes
Oe2/71	307/307 (BB) 307/309 (AB) 309/309 (AA)	12 31 5	6.125 (*)	AB: BB	31: 12	0.194 (NS)	Deficiency in homozygote IBD (AA) (L002-55)
Oe3/44	211/211 (BB) 202/211 (AB) 202/202 (AA)	2 27 19	12.792 (***)	AB: AA	27: 19	6.522 (*)	Significant excess of (AA)
HA1	151/151 (BB) 132/151 (AB) 132/132 (AA)	13 30 4	7.043 (**)	AB: BB	30: 13	0.628 (NS)	Deficiency in homozygote IBD (AA) (L002-55)
HA7	156/156 (BB) 156/184 (AB) 184/184 (AA)	7 33 8	6.792 (**)	AB: BB AB: AA	33: 7 33: 8	1.2 (NS) 0.658 (NS)	Deficiency in homozygote IBD (AA) and (BB)
HA21	158/158 (BB) 158/180 (AB) 180/180 (AA)	3 39 3	24.2 (***)	AB: BB AB: AA	39: 3 39: 3	7.143 (**) 7.143 (**)	Significant excess of heterozygotes
B11	163/163 (BB) 116/163 (AB) 119/119 (AA)	6 35 7	10.125 (**)	AB: BB AB: AA	35: 6 35: 7	2.349 (NS) 1.556 (NS)	Deficiency in homozygote IBD (AA) and (BB)
C6	104/104 (BB) 104/106 (AB) 106/106 (AA)	14 31 3	9.125 (**)	AB: BB	31: 14	0.896 (NS)	Deficiency in homozygote IBD (AA) (L002-55)

For OE.F2.04.63, three types of segregation could be observed for the microsatellites, AA: AB (1: 1), AA: AB: AC: BC (1: 1: 1: 1) or AC: AD: BC: BD (1: 1: 1: 1). Seven markers showed significant departures from Mendelian segregation that was due to a deficiency in homozygote IBD (AA): J12, U2, T5, Oe1/63, Oe3/44, HA1 and C6. For Oe3/44, the AA homozygote IBD was completely absent. Four of these seven markers, J12, U2, HA1 and C6, were not significantly distorted when a χ^2 goodness of fit AB: AC: BC (1: 1: 1) was applied. The significant Mendelian segregation test AA: AB: AC: BC was therefore mainly due to a deficit in the IBD homozygote for these 4 microsatellites. Four other markers, Oe2/71, O9, G9 and HA7 exhibited segregation distortion that was not due to a deficit in IBD homozygote (Figure 89).

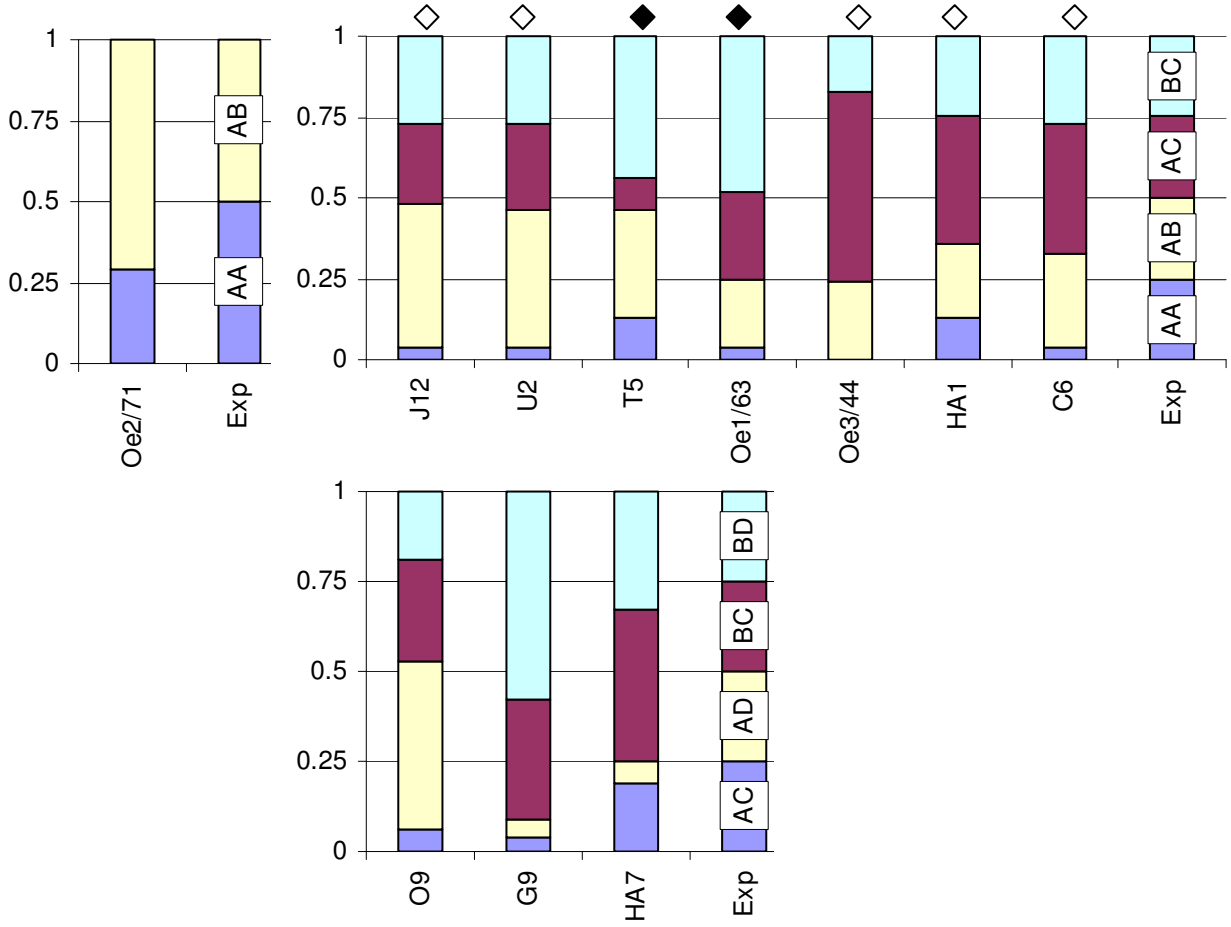


Figure 89. Relative proportions of microsatellite genotypes in 11 cases (out of 17) for which there was a significant departure (p<0.05) from expected Mendelian segregation ratios for the family OE.F2.04.63. Three types of segregation are shown, as indicated by the far-right bar of each type labelled "Exp" which shows the expected Mendelian proportions of each genotype. Open diamonds denote markers for which the AA class is homozygote IBD for a L002-53 grand-parental allele and closed diamonds for a W102 grand-parental allele.

III.2. Comparison of heterozygosity between the families OE.F2.04.63 and OE.F2.04.45

Observed heterozygosity (H_0) and expected heterozygosity (H_{nb}) are shown for the 2 families of increasing inbreeding: OE.F2.04.63 (f=0.25) and OE.F2.04.45 (f=0.50) in Table 37. For the family OE.F2.04.63, expected heterozygosity was either 1.00 (type 1:1:1:1 with 4 segregating alleles), 0.75 (type 1:1:1:1 with 3 segregating alleles), 0.50 (type 1:2:1 or type 1:1) or 0.00 (allele fixed). For each marker, the observed heterozygosity was equal (for markers with an H_{nb} of 1.00) or higher than the expected heterozygosity. However, a χ^2 goodness of fit test performed on the 17 microsatellites (with H_{nb} >0) was not significant: the observed heterozygote excess was not significant (χ^2=24.48, 16 ddf, p>0.05). For the family OE.F2.04.45, due to the self-fertilisation of the F1 parent, expected heterozygosity was 0.50 for all markers (parent was heterozygous) except the ones that were fixed for an allele (parent was homozygous, therefore H_{nb}=0.00). For each marker, the observed heterozygosity was higher than the expected one. Overall, a χ^2 goodness of fit test performed on the 15 microsatellites (with H_{nb} >0) revealed that the observed heterozygote excess was highly significant (χ^2=69.20, 14 ddf, p<0.001) (Table 37).

Table 37. Observed heterozygosity (H_0) and expected heterozygosity (H_{nb}) for families OE.F2.04.63 and OE.F2.04.45. N: number of F2 progeny in each family; f: inbreeding coefficient; NS: χ^2 goodness of fit test not significant; p<0.001: χ^2 goodness of fit test significant at α=1‰.

Locus		OE.F2.04.63 N=48 (f=0.25)	OE.F2.04.45 N=48 (f=0.50)	Additional informations
*Oedu*J12	H_{nb}	0.75	0.50	
	H_0	0.96	0.73	
*Oedu*U2	H_{nb}	0.75	0.50	
	H_0	0.96	0.69	
*Oedu*H15	H_{nb}	0.00	-	Locus fixed for one allele for family 04.63.
	H_0	0.00	-	Null allele segregating in family 04.45: missing genotypes
*Oedu*T5	H_{nb}	0.75	0.50	
	H_0	0.88	0.81	
*Oedu*O9	H_{nb}	1.00	0.50	
	H_0	1.00	0.69	
*Oedu*T18	H_{nb}	0.50	-	Null allele segregating in family 04.45.
	H_0	0.63	-	Null allele segregating in family 04.45: missing genotypes
*Oedu*G9	H_{nb}	1.00	0.50	
	H_0	1.00	0.63	
Oe1/47	H_{nb}	0.75	0.50	
	H_0	0.66	0.77	
Oe1/63	H_{nb}	0.75	0.50	Locus fixed for one allele for family OELL305.
	H_0	0.96	0.81	

Oe1/64	H_{nb}	0.00	-	Locus fixed for one allele for family 04.63. Nul allele segregating in family 04.45: missing genotypes
	H_0	0.00	-	
Oe2/71	H_{nb}	0.50	0.50	
	H_0	0.71	0.65	
Oe3/37	H_{nb}	0.50	0.00	Locus fixed for one allele for family 04.45.
	H_0	0.58	0.00	
Oe3/44	H_{nb}	0.75	0.50	
	H_0	1.00	0.56	
Oedu.HA1	H_{nb}	0.75	0.50	
	H_0	0.88	0.64	
Oedu.HA7	H_{nb}	1.00	0.50	
	H_0	1.00	0.69	
Oedu.HA11a	H_{nb}	1.00	0.50	
	H_0	1.00	0.64	
Oedu.HA21	H_{nb}	0.50	0.50	
	H_0	0.52	0.87	
Oedu.B11	H_{nb}	1.00	0.50	Null allele segregating in family 04.63 but never at the homozygous state so no missing genotypes.
	H_0	1.00	0.73	
Oedu.C6	H_{nb}	0.75	0.50	
	H_0	0.96	0.65	
Total	H_{nb}	0.68	0.47	
	H_0	0.77	0.66	
Overall test	χ^2 *value*	24.48, 16 ddf, NS	69.20, 14 ddf, $p<0.001$	Strong heterozygote excess for the family OE.F2.04.45

III.3. Investigation into the growth/heterozygosity relationship in the family OE.F2.04.45

The number of microsatellite loci at which individuals were heterozygous ranged from 5 to 14. Growth rate for total length (final length minus initial length, divided by initial length) ranged from 13% to 127% and growth rate for weight data (final weight minus initial weight, divided by initial weight) from 139% to 478%. For a particular number of heterozygous markers, there was a wide range of growth rate. No relationship could be found between the number of heterozygous microsatellite markers and the growth rate computed from total length or weight data (Figure 90).

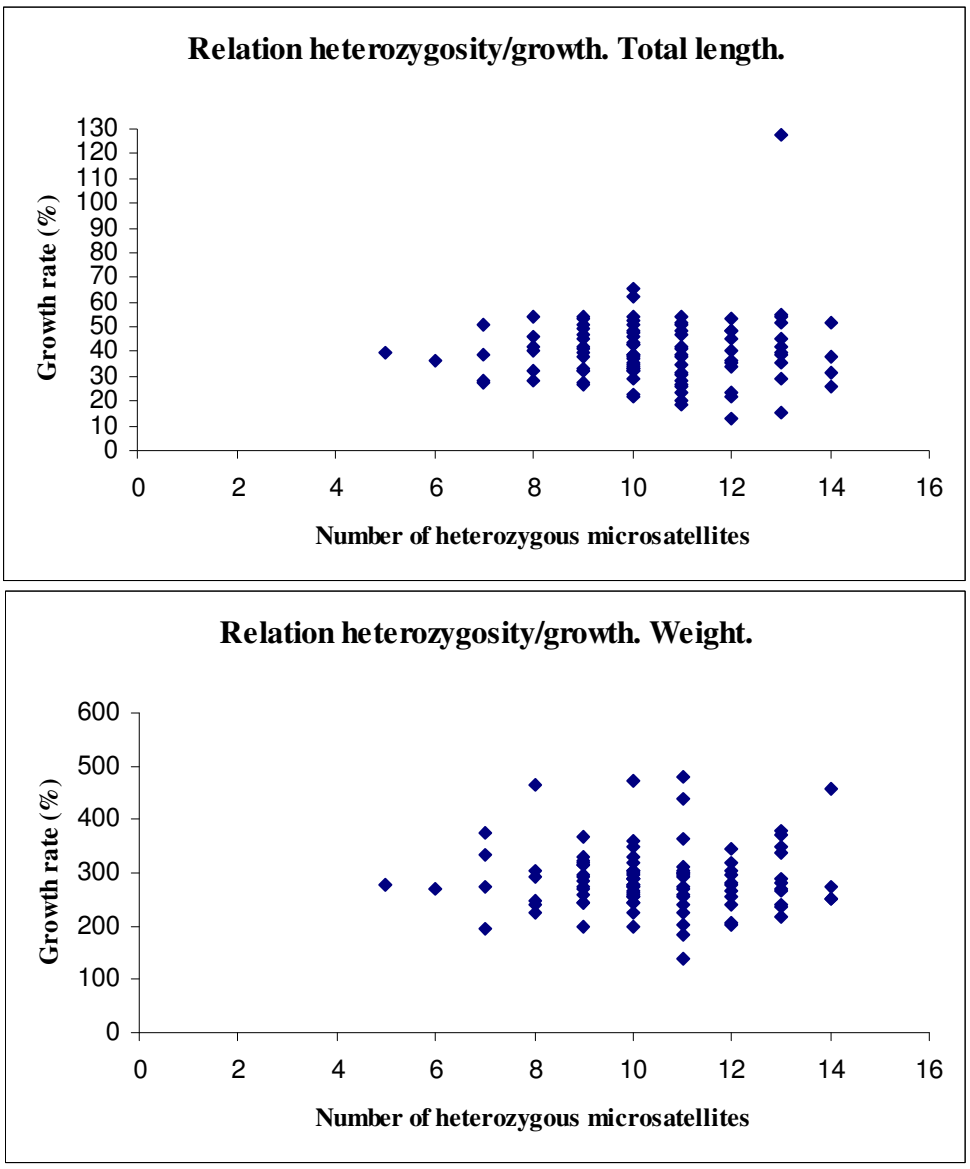


Figure 90. Relationship between the number of heterozygous microsatellites and the growth rate (%), for total length data or weight data in the family OE.F2.04.45. Growth rates were computed by (final-initial)/initial x 100.

Growth rate was computed marker by marker, for each marker genotype. As family OE.F2.04.45 issued from the self-fertilisation of an F1 individual, for each marker 3 genotypic classes could be found, classes 1 and 3 being the IBD homozygotes and class 2 being the heterozygote genotype. As stated earlier, the only exception concerned marker Oe1/47 for which the IBD status could not be determined. For all markers, no relationship was found between the marker genotype and the growth rate computed from total length data (Figure 91) or from weight data (Figure 92).

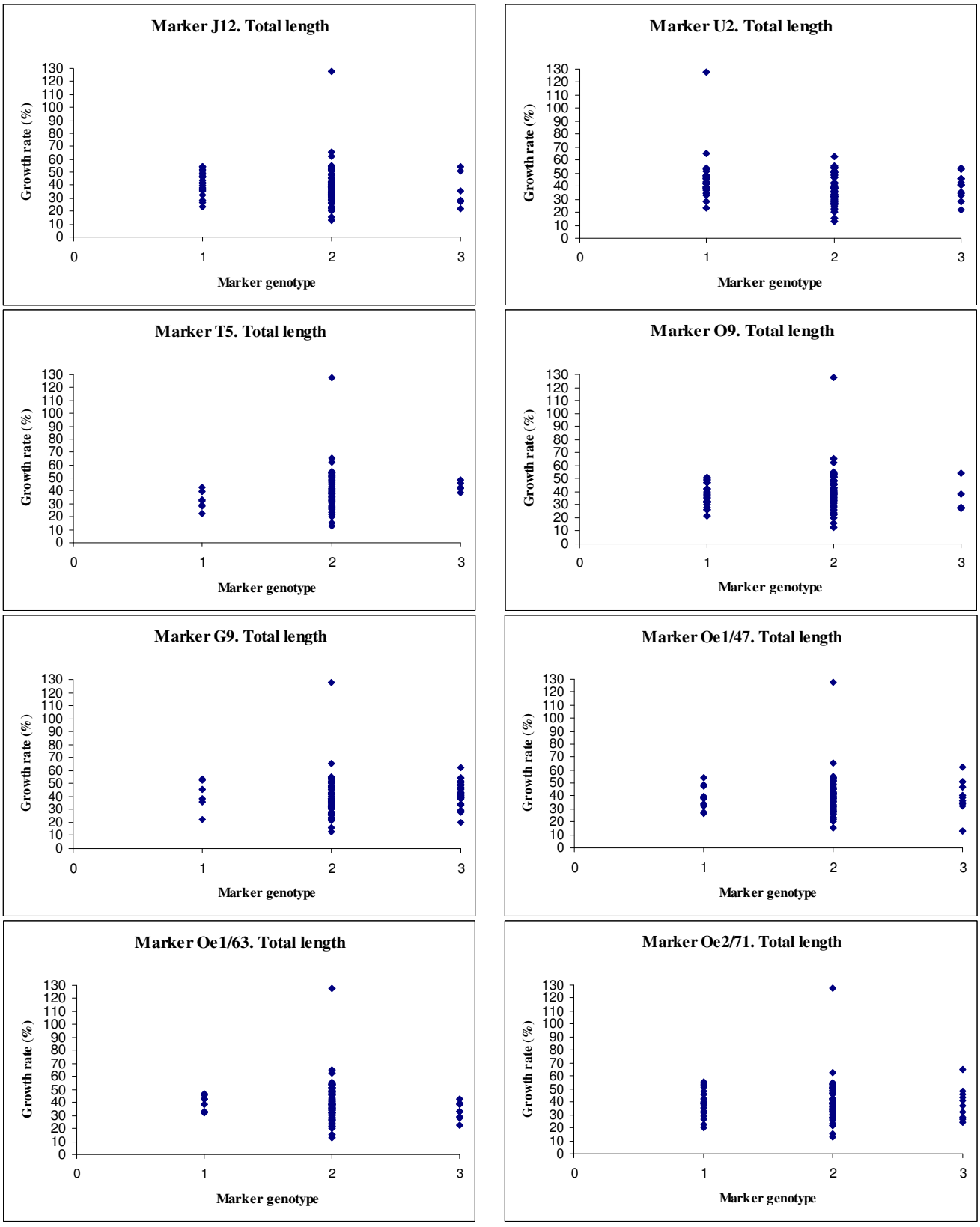


Figure 91-continued

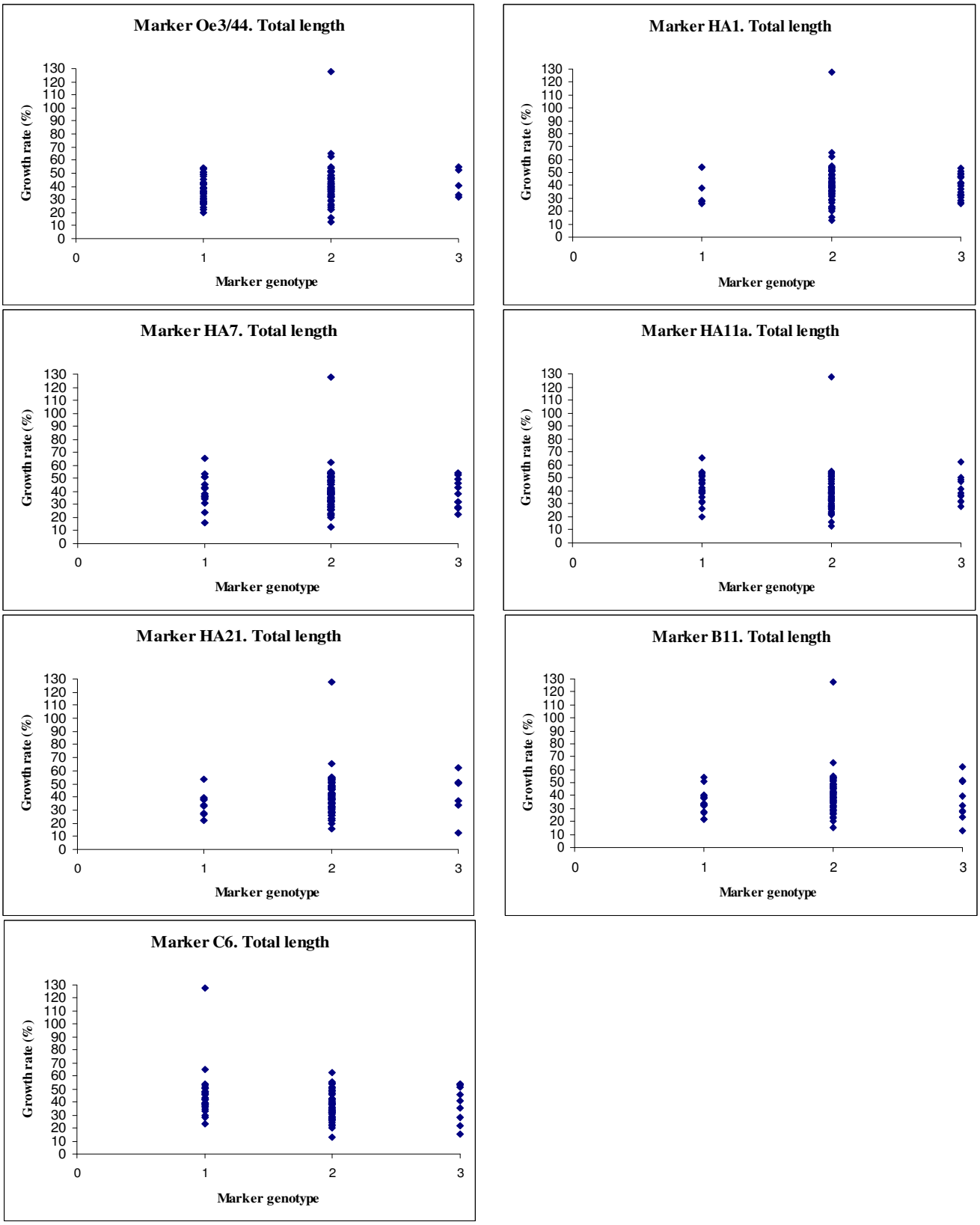
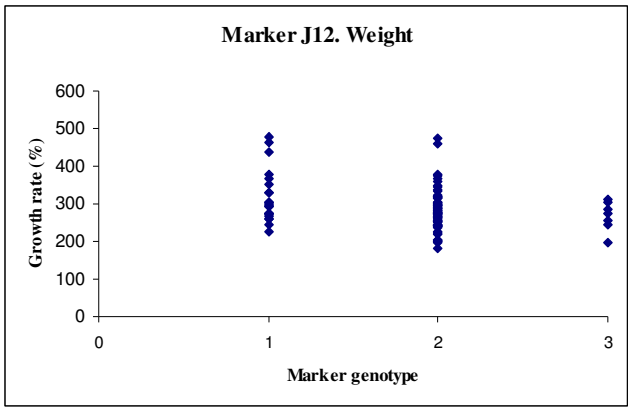
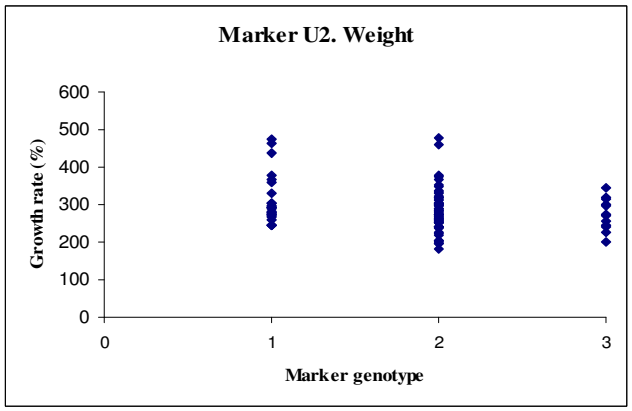
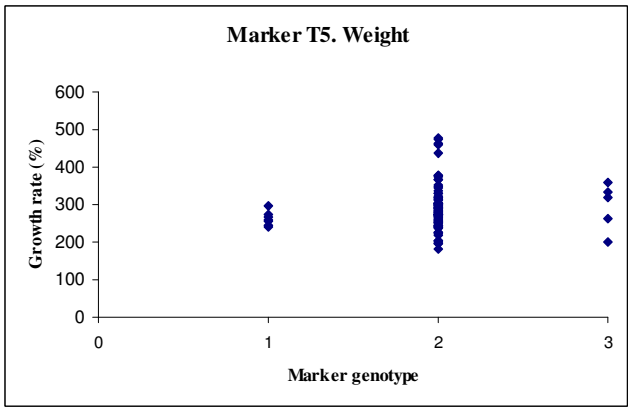
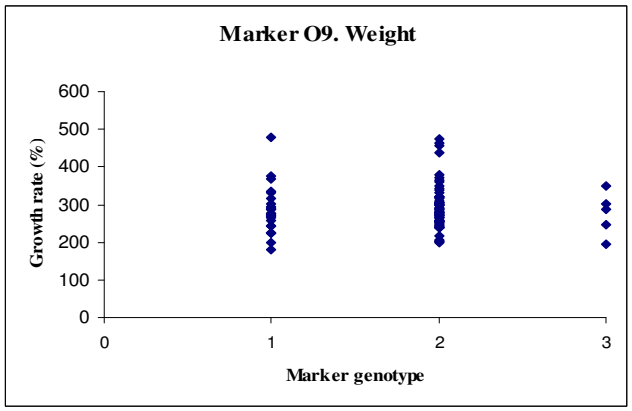


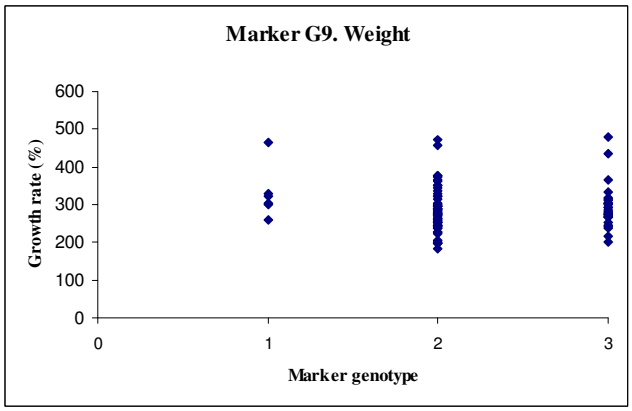
Figure 91. Relationship between microsatellite genotype and the growth rate (%) computed from total length data in the family OE.F2.04.45. Classes 1 and 3: IBD homozygotes; class 2: heterozygote.

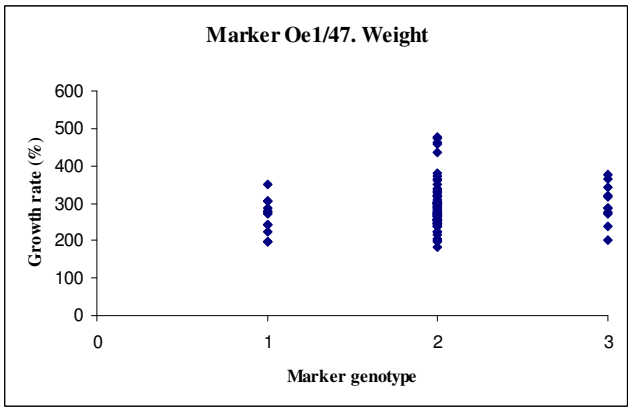


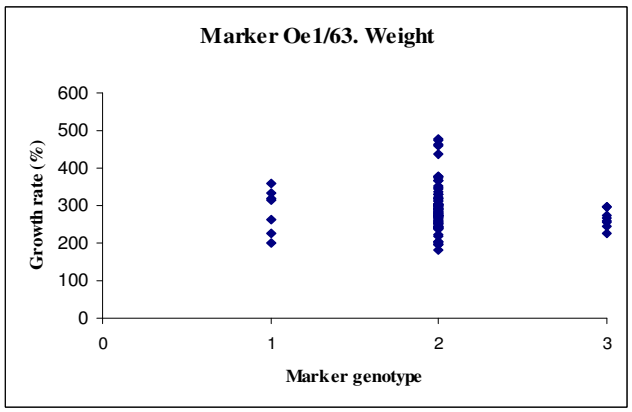












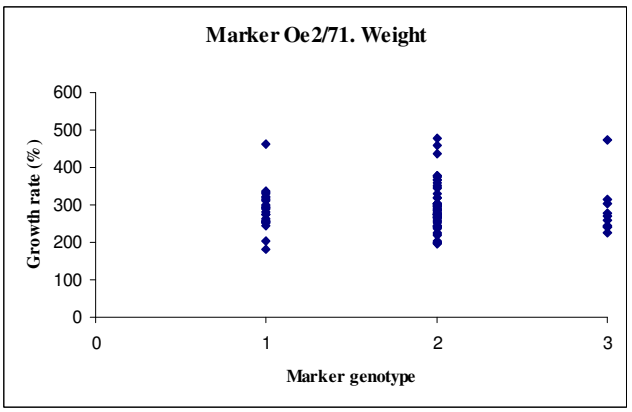


Figure 92-continued

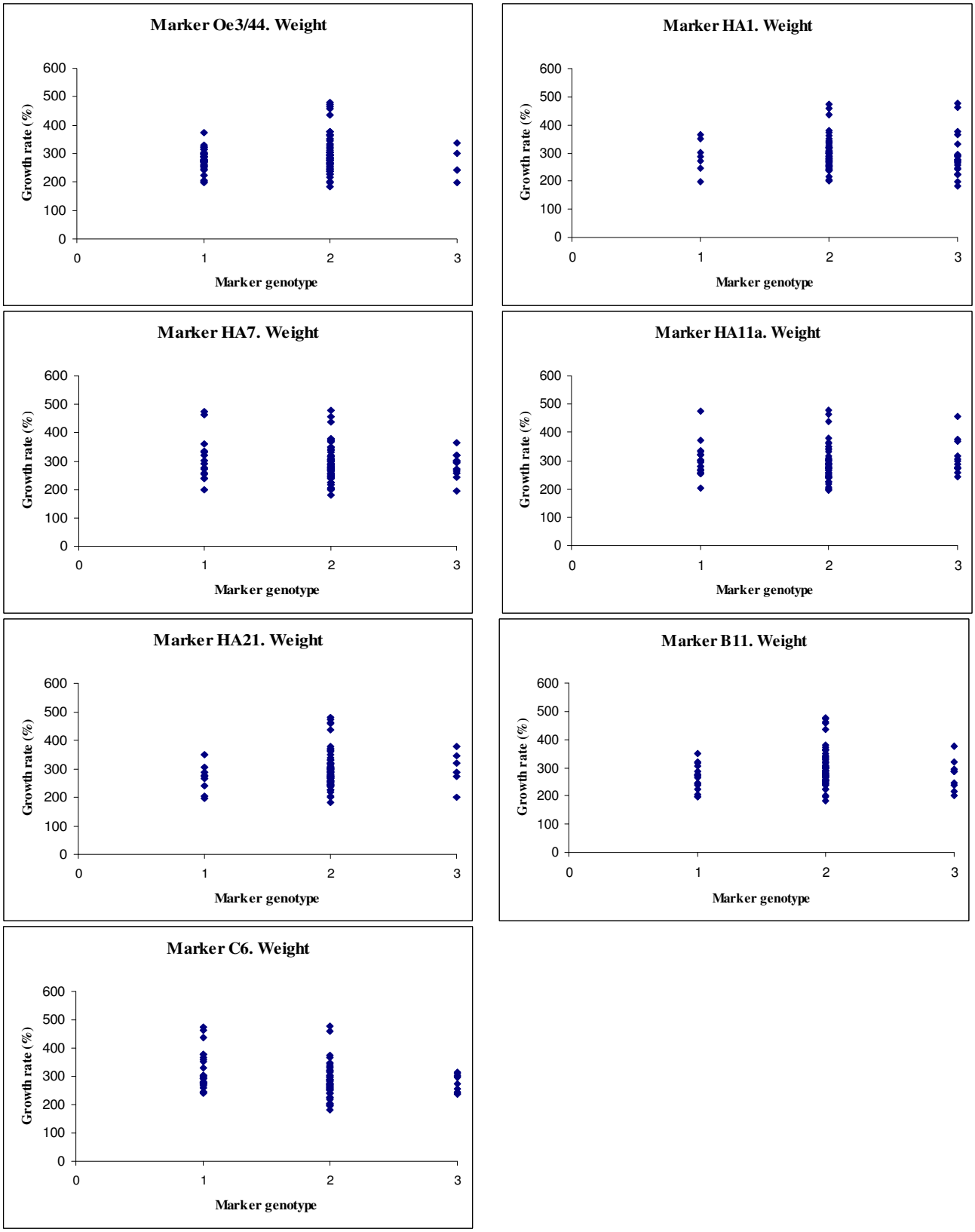


Figure 92. Relationship between microsatellite genotype and the growth rate (%) computed from weight data in the family OE.F2.04.45. Classes 1 and 3: IBD homozygotes; class 2: heterozygote.

IV- DISCUSSION

I.V.1. High genetic load in *O. edulis*

Relatively high segregation distortion was reported in the two segregating families of *O. edulis* analysed: 51.1% of the overall markers for the family OE.F2.04.45 (n=48) and 31.9% of overall markers for the family OE.F2.04.63 (n=48) (Figure 87). Moreover, the family derived from the self-fertilisation of an F1 parent (OE.F2.04.45) exhibited significantly higher segregation distortion for the AFLP markers ($p<0.001$) and overall markers (microsatellites and AFLPs) ($p<0.001$).

The two mapping families (OE.F2.04.45 and OE.F2.04.63) originated from a 6th generation inbred line (6 generations of full-sib matings) that would certainly have undergone some inbreeding depression. Therefore, assuming that purging of deleterious genes by full-sib crosses was not complete by the 6th generation, it is likely that the high segregation distortion observed was due to linkage of markers with lethal or deleterious genes in the recessive state. This high genetic load has previously been reported in *O. edulis* (Bierne et al., 1998) and in other bivalve species such as *C. gigas* (McGoldrick and Hedgecock, 1997; Launey and Hedgecock, 2001). It is moreover interesting to note that level of segregation distortion was significantly higher in the family derived from a self-fertilisation (f=0.50), and hence in the family with the highest level of inbreeding. The link between the level of segregation distortion and the level of inbreeding coefficient reinforced our hypothesis that the high distortion in segregation patterns observed was probably due to a high genetic load.

The vast majority of significant segregation distortions of microsatellites resulted from a deficiency of an IBD (identical by descent) homozygous genotype (7/11 for OE.F2.04.63 and 14/15 for OE.F2.04.45) (Figures 88 and 89). It was interesting that most alleles that were associated with a deficiency in an IBD homozygous genotype came from the grand-parent derived from the inbred line (L002-55 or L002-53). Indeed, for OE.F2.04.45, 13 out of 14 cases showed deficiencies of the IBD homozygote AA (L002-55 grand-parental allele). For OE.F2.04.63, 5 out of 7 cases showed deficiencies of the IBD homozygotes for a L002-53 grand-

parental allele. Therefore, the majority of significant distortions could be attributed to selection against recessive deleterious mutations at linked fitness genes.

Moreover, an excess of heterozygotes could be observed in the two families. However, this observed heterozygote excess was significant only in the self-fertilised family, whereas in the family OE.F2.04.63 this excess was just above the significance threshold (Table 37). Therefore, the expression of deleterious genes was stronger in the family with the highest level of inbreeding, corroborating the previous findings concerning the link between genetic load and level of inbreeding.

Finally, because of this high genetic load in *O. edulis*, the production of homozygous inbred lines in that species could be difficult to achieve. Due to the expression of deleterious genes, it is expected that only heterozygotes survive; the deleterious allele is maintained at the heterozygote state. Therefore, in the future production of homozygous lines, the gain in homozygosity will be possible only after purging those deleterious mutations. Some evidence of the possibility of purging the genetic load was reported based on experimental data (Crnokrak and Barrett, 2002) or on simulation studies (Fu, 1999). However, mechanisms involved in the purging of deleterious genes by inbreeding are still poorly understood and under debate (Fu, 1999). The effectiveness of purging would depend on the level of the deleterious effects (lethal or sub-lethal genes being more easily purged) and increase with high dominance, strong synergism and low inbreeding (Fu, 1999) but stochastic processes would play an important role (Miller and Hedrick, 2001). Several studies agreed that purging the genetic load by inbreeding may not be a good strategy (Fu, 1999; Miller and Hedrick, 2001). In our study, it was interesting that almost all the deficiencies in an IBD homozygous genotype corresponded to an allele coming from the grand-parent derived from the inbred line. This result was consistent with the absence of purging in a 6^{th} generation inbred line and therefore strengthened the finding of Fu (1999) and Miller and Hedrick (2001) that successive inbreeding may not be a powerful strategy for the purging of the genetic load. Based on the potential high genetic load in *O. edulis*, diploid gynogenesis would probably represent a promising (and alternative) strategy for the production of homozygous inbred lines. Such strategy has been used for genetic and QTL mapping in several fish species, such as rainbow trout *Oncorhynchus mykiss* (e.g. Nichols et al., 2003; Zimmerman et al., 2005)

or zebrafish *Danio rerio* (e.g. Kelly et al., 2000). However, diploid gynogenesis in oysters is difficult to achieve and brooding oysters such as *O. edulis* are even less amenable to ploidy manipulation (Beaumont and Fairbrother, 1991).

IV.2. MLH/growth relationship

Weak, but often significant correlations between multiple locus heterozygosity at allozyme or microsatellite loci and fitness-related characters (heterozygosity- fitness correlations, HFC) have been detected in oysters and other bivalves (Alvarez et al., 1989; Zouros and Pogson, 1994; David et al., 1995; David, 1998; Bierne et al., 1998; Launey and Hedgecock, 2001). HFC can be summarised as meaning that individuals which are more heterozygous, on average, grow faster, have a lower standard respiration rate, or a greater efficiency for protein synthesis and a higher scope for growth than individuals which are more homozygous. Although the use of different terminology by different authors has caused some confusion, HFC is thought to be a result of either (a) "direct overdominance" at the scored loci and/or (b) "associative overdominance" the association between homozygosity at the scored loci and homozygosity for deleterious genes at loci in linkage disequilibrium with the scored loci (David, 1998). Direct overdominance is usually the genetic explanation given for the phenotypic phenomenon of "heterosis" while associative overdominance is the classic explanation for the phenotypic phenomenon of inbreeding depression. However, it was surprising to discover that HFCs could be detected when only considering a very small part of the genome (a few allozyme or microsatellite loci). In fact, estimations of the number of lethal genes present in the oyster genome (Bierne et al., 1998; Launey and Hedgecock, 2001) have demonstrated the potential high genetic load in bivalves and this would make detection of HFC more likely.

If the HFC is due to general exposure of deleterious recessives across the genome (associative overdominance) then we might not expect to find that distorted markers cluster in specific regions of the genome (spanning small genetic distances) as has been shown in this study for *O. edulis* and also in *C. virginica* and *C. gigas* (Yu and Guo, 2003; Li and Guo, 2004).

No significant HFC (for growth) was evident in the family of *O. edulis* studied here (Figures 90, 91, 92) but this is not unexpected because HFC is not usually detected in data sets from a restricted genetic background (Beaumont, 1991). Also, microsatellites are supposedly neutral and this would make HFCs less likely.

IV.3. Future work

The effect and the number of lethal genes should be estimated by applying a two-locus selection model, based on the deficiencies of IBD homozygotes, as described in Launey and Hedgecock (2001) for *C. gigas*. This would allow the estimation of two parameters which could account for the observed departures from Mendelian segregation: the selection coefficient of the deleterious gene and the map distance between it and the marker. Moreover, the comparison of the number of lethal genes in the two different families would be interesting to see if those estimations are in agreement in the two families.

CHAPTER 7

General discussion and conclusions

Genetic linkage maps consist of ordering molecular markers across the genome and require a high number of markers for a good coverage of the genome. Such maps represent a framework which enables the identification and localisation of QTLs for traits of interest, such as growth or disease resistance, with the final aim of achieving genetic improvement through marker-assisted selection (Liu and Cordes, 2004). Therefore, they represent new genomic tools that will assist any selective breeding programme. Data on bivalves are scarce. Genetic linkage maps have been established in several commercially important bivalves: the Pacific oyster *Crassostrea gigas* (Hubert and Hedgecock, 2004; Li and Guo, 2004), the Eastern oyster *Crassostrea virginica* (Yu and Guo, 2003) and the Zhikong scallop *Chlamys farreri* (Wang et al., 2004; Li et al., 2005; Wang et al., 2005). No genetic map has yet been constructed in any mussel or flat oyster species.

The flat oyster industry has endured a drastic decline in its production, due to two parasitic diseases: marteiliasis (due to *Marteilia refringens*) and bonamiasis (due to *Bonamia ostreae*). Since 1985, Ifremer has been undertaking a selective breeding programme for resistance to bonamiasis that led to the production of two improved oyster strains. These selected oyster strains represent a valuable resource in the context of QTL mapping.

Therefore, the aims of this PhD thesis were to:

- establish a first genetic linkage map in the blue mussel *Mytilus edulis*
- establish a first genetic linkage map in the European flat oyster *Ostrea edulis*
- find QTLs of resistance or susceptibility to bonamiasis in *O. edulis*, based on a 6-month trial challenge experiment by cohabitation of wild oysters (overinfected with *B. ostreae*) and tested oysters (segregating family)

I- AFLP METHODOLOGY

Only a few microsatellites have been published so far for the blue mussel *M. edulis* (Presa et al., 2002) although there are more for the European flat oyster *O. edulis* (Naciri et al., 1995; Launey, 1998; Morgan et al., 2000; Morgan and Rogers, 2001; Sobolewska et al., 2001; Launey et al., 2002).

Due to the relatively limited number of microsatellites available in the two species studied, the AFLP methodology (Vos et al., 1995) was chosen for developing molecular markers in the two species. The AFLP methodology is very advantageous in species with poor genetic resources because they allow the generation of a high number of molecular markers relatively quickly (several markers can be detected in a single PCR assay), require no prior knowledge of the genome and the methodology is easy to transfer from one species to another. Indeed, the same protocol was used in *M. edulis* and *O. edulis* and demonstrated that the AFLP methodology is an efficient way of generating a high number of molecular markers that can be used for genetic mapping purposes. The overall rate of polymorphism reported in this study in the blue mussel was 33.6%, averaging 22 markers generated per primer pair. Therefore, this methodology was critical to the success of the genetic mapping projects. However, AFLP markers are dominant: their genotypes need to be inferred from phenotypes (band absence or presence). The band absence phenotype is associated to the homozygous recessive genotype *aa*, whereas the band present genotype could be associated to two genotypes, *AA* or *Aa*. Because genetic mapping relies on the estimation of recombination frequencies between pairs of markers and implies to be able to distinguish parental from recombinant gametes, missing genotypes (not distinguishing *AA* from *Aa*, referring to *A?* genotype instead) would hamper the good success of establishing a genetic map. Therefore, the dominance feature of the AFLPs required the use of suitable mapping family and experimental design to limit the number of missing genotypes (*A?*).

II- MAPPING FAMILY AND EXPERIMENTAL DESIGN

In the flat oyster *O. edulis*, the mapping families consisted of a three-generation pedigree (grand-parents, parents and offspring) that did not come from truly inbred lines (homozygous for all loci), but from the cross between a 6^{th} generation inbred line and a wild oyster (families OE.F2.04.45 and OE.F2.04.63) or from an oyster of a selected family and a wild oyster (family OE.F2.05.04). This mating scheme was unusual for a mapping family in a shellfish species. Indeed, the classical mating schemes in experimental populations where inbred lines are available generally involve the analysis of either backcross, or F2 progeny. Some mapping panels reported in the literature consisted of a three-generation pedigree of backcross families (Naruse et al.,

2000; Sakamoto et al., 2000; Li and Guo, 2004), or F2 families (Shimoda et al., 1999; Li et al., 2003). Other mapping panels consisted of two generation pedigree (parents and offspring) (Coimbra et al., 2003). However, when studying natural populations, or when inbred lines are not available, individuals can be taken from the population, genotyped and mated in pairs to yield a number of full-sib families. In a particular family, any pair of segregating loci will represent either an F2 (if both parents are heterozygous for the pair of markers) or a backcross (if only one parent is heterozygous whereas the other is homozygous) (Kearsey and Pooni, 1998). This strategy has been used in several studies (Waldbieser et al., 2001; Wilson et al., 2002; Yu and Guo, 2003; Hubert and Hedgecock, 2004) and is the most commonly used experimental design in shellfish species.

In the context of QTL mapping, a three-generation pedigree was chosen for the mapping family in *O. edulis* because it allows the estimation of the recombination fraction between the molecular marker and the QTL as well as the estimation of both the additive and dominant effects of the QTL (see Chapter 1). This experimental design proved to be efficient for the mapping of microsatellites (87.5% mapped in both P1 and P2 parents of the family OE.F2.04.63; 94.1% and 77.8% mapped in the parents 410_7 and 410_8 of the family OE.F2.05.04) and type I AFLPs (87.3% mapped in P1, 90.5% mapped in P2, 83.9% mapped in 410_7 and 89.9% mapped in 410_8). However, as expected according to the informativeness of the markers (Ritter et al., 1990), the mapping of type III AFLPs (3:1 segregation type) in a F2-type family was less powerful because only 32.3% and 33.8% were mapped in the two P1 and P2 parental maps of the family OE.F2.04.63; 37.8% and 28.9% in the 410_7 and 410_8 parental maps of the family OE.F2.05.04. The low information content of type 3:1 AFLP markers and therefore the difficulty in accurately estimating recombination frequency between two 3:1 AFLPs was confirmed in the family OE.F2.04.45. Indeed, due to the family structure (self-fertilisation of an F1 parent), all AFLP markers were of 3:1 segregation type. This mapping family proved to be unsuitable for genetic mapping. The 3:1 AFLPs were highly clustered because recombination could not be identified between them; only recombination frequency between a microsatellite marker and a 3:1 AFLP or between 2 microsatellites could be accurately estimated, each microsatellite being associated with a cluster of 3:1 AFLPs (Figures 54 and 55). In conclusion, the main problem of using an F2-type mapping family with dominant markers (AFLPs in our study) comes from the

missing genotypes (band presence related to a A? genotype). In a three-generation pedigree, only AFLPs that are present in only one grand-parent and only one F1 parent are fully informative. So, overall, only a small proportion of the high number of markers generated can effectively be used for genetic mapping.

Due to the loss of informativeness of AFLPs in a F2-type mating sheme, and because QTL mapping was not an aim in the mussel, a full-sibs family (consisted of 2 parents and their offspring) was chosen for the genetic mapping project of the blue mussel *M. edulis* and a pseudo-test cross strategy applied (Grattapaglia and Sederoff, 1994). This mapping strategy relies on the selection of single-dose polymorphic markers that are heterozygous in one parent (*Aa*) and recessive homozygote (*aa*) in the other, and are expected to segregate 1:1 in their progeny. This strategy, coupled with the AFLP technology (generation of a high number of markers relatively quickly and at a reasonable cost), has been employed in a variety of species, from shellfish (Li et al., 2003; Yu and Guo, 2003; Li and Guo, 2004; Perez et al., 2004) to plant species (Verhaegen and Plomion, 1996; Barreneche et al., 1998). The establishment of a genetic linkage map in the blue mussel *M. edulis* has confirmed that this approach is a powerful one for the establishment of genetic linkage maps.

III- GENETIC MAPPING

Three genetic mapping softwares were used during this project: MapMaker (Lander et al., 1987) and JoinMap (Stam, 1993) for the mussel genetic mapping project, and CriMap (Green et al., 1990) for the flat oyster mapping project. All three softwares use the maximum likelihood approach to group markers into linkage groups and estimate the recombination frequencies between them. The main differences between them concern the coding of the data, the interface under which the software is running and more importantly the genotypic data that each software can handle.

CriMap software required no particular coding of the data, each allele being input in a column, and the format of the data is very flexible (letters, arbitrary numbers: 1, 2, 3, 4 or numbers that refer to the fragment size in base pair). The input file for CriMap software was

therefore relatively easy to create; the "prepare" command transforming the raw data into a usable format (phase inferred by the software). However, CriMap software runs in a UNIX environment and required the typing in of all the commands, making analyses time-consuming and very repetitive. In addition, identifying linkage groups under CriMap has to be done by hand based on the two-point data (LOD score and recombination frequency for each pair of marker asscociated with a LOD score above a threshold set by the user). Finally, CriMap is able to analyse any type of data and family structure, probably linked to the simplicity of the input data file (no particular coding). It proved to be suitable for our *O. edulis* dataset: F2-type family, with a mixture of 1:1 AFLPs, 3:1 AFLPs and microsatellite markers.

MapMaker and JoinMap software both required a particular coding of the data. This coding was initially daunting, each genotype being associated with a letter depending on the genotypes of the progeny and their parents. MapMaker runs under a DOS or UNIX environment, each command being typed in by the user. In contrast, JoinMap runs under a Windows environment and the interface was far more user-friendly and the data more easy to manipulate and experiment with. Indeed, with JoinMap, it was easy to include or discard markers and/or individuals from the analyses, by ticking or un-ticking boxes. However, the identification of linkage groups was easier with MapMaker because two parameters (maximum recombination frequency and LOD score) can be used simultaneously to find the linkage groups. In contrast, with JoinMap, only one parameter could be used at a time. However, MapMaker was not able to handle 1:1 and 3:1 markers in the same dataset, contrary to JoinMap. Therefore, in the mussel genetic mapping project, the combination of the two softwares, MapMaker and JoinMap, was very useful. MapMaker was first used for finding the linkage groups and establishing sex-specific framework groups and then JoinMap was used to establish a consensus map based on the 3:1 markers and to map the associated markers.

The programme of study led to the establishment of first genetic linkage maps in two bivalve species of economical and scientific importance, the blue mussel *M. edulis* and the flat oyster *O. edulis*. The maps established, even if preliminary, achieved a relatively good genome coverage, above 80%, that was in the range of the ones found in the literature. Some common features between the maps established in the two species could be highlighted:

- the number of linkage groups was close or equal to the haploid number of chromosomes reported in the two species (10 for *O. edulis* and 14 for *M. edulis*)
- clustering of AFLPs was observed
- gaps remained to be filled (large intervals with a genetic distance above 20 cM were observed in the two species)
- some small linkage groups (covering a distance of less than 20 cM, with a few markers) were observed

Whether the observation of large intervals and small linkage groups reflect true patterns of differences in recombination rates according to the regions of the genome (hotspot of recombination, versus suppression of recombination), or artefacts of genetic mapping remain to be seen. Indeed, several linkage groups may in fact correspond to the same chromosome and one can imagine that adding more markers would merge these small groups into a bigger one and that gaps would be filled.

In order to evaluate the robustness of the linkage maps established in *O. edulis*, it was interesting to compare the order of markers obtained in two different mapping families, OE.F2.04.63 and OE.F2.05.04. Thanks to the segregation of microsatellite markers, 5 homology groups could be found between the two mapping families. Homology group 1 was based on the co-segregation of 4 microsatellites (*Oedu*.B11, *Oedu*.HA21, Oe1/47 and Oe3/37). Moreover, 3 AFLPs were mapped in the two mapping families, A3f73 and E1f310 in G3_410_8 and G3_P2, E4f291 in G1_410_7 and G3_P2 (Figure 93). Homology group 2 was based on the co-segregation of 5 microsatellites (*Oedu*U2, *Oedu*C6, *Oedu*J12, *Oedu*O9 and *Oedu*.HA1). The AFLP marker D1f203 was mapped in G2_410_7, G2_410_8 and G4_P2 (Figure 94). Homology group 3 was based on the mapping of 2 microsatellites (*Oedu*.HA11 and Oe2/71) that were both mapped in G5_P1. Therefore homology between G5_410_7 and G7_410_8 was probable. Moreover, the AFLP B4f266 was mapped in G5_410_7 and G6_P2; and two pairs of AFLPs that differed from one another by 1 bp reinforced the homology: A4f313 and A11f179 (mapped in terminal position of G7_410_8) and A4f312 and A11f178 (mapped in terminal position of G5_P1) corresponded probably to the same molecular markers (Figure 95). Homology group 4 was based on the co-segregation of 3 microsatellites (*Oedu*.HA7, *Oedu*T5 and Oe1/63). Moreover, 3 AFLPs were mapped in the two mapping families, E9f368 and A10f137 in

G6_410_8 and G6_P1; B12f258 in G7_410_7 and B12f257 in G2_P2 (Figure 96). Homology group 5 was based on one microsatellite (Oe3/44) (Figure 97). For the first 4 homology groups, several markers were mapped in the two mapping families and marker orders were respected in most cases. The main discrepancy occurred for the group G4_P2 where microsatellites were mapped in different orders from the 3 other parental groups, G2_410_7, G2_410_8 and G1_P1 (Figure 94).

Despite the lack of anchor loci (microsatellites), these results are encouraging because 5 homology groups could be identified between two different mapping families and most markers orders were respected. It is moreover interesting to note that a few AFLPs were mapped in the two mapping families, addressing therefore the possibility of the transferability of the AFLP markers from one family to another. Because the AFLP methodology relies on the accurate estimation of the length of restriction fragments, transferability of AFLP markers would be possible only if the genotyping is made with the same scoring system (acrylamide gel, or capillary array). Systems based on capillary array will probably be more suitable for transferring AFLP markers as long as room temperature is well-controlled because such systems are less dependant on user-variability (occurring mainly during the gel polymerisation process).

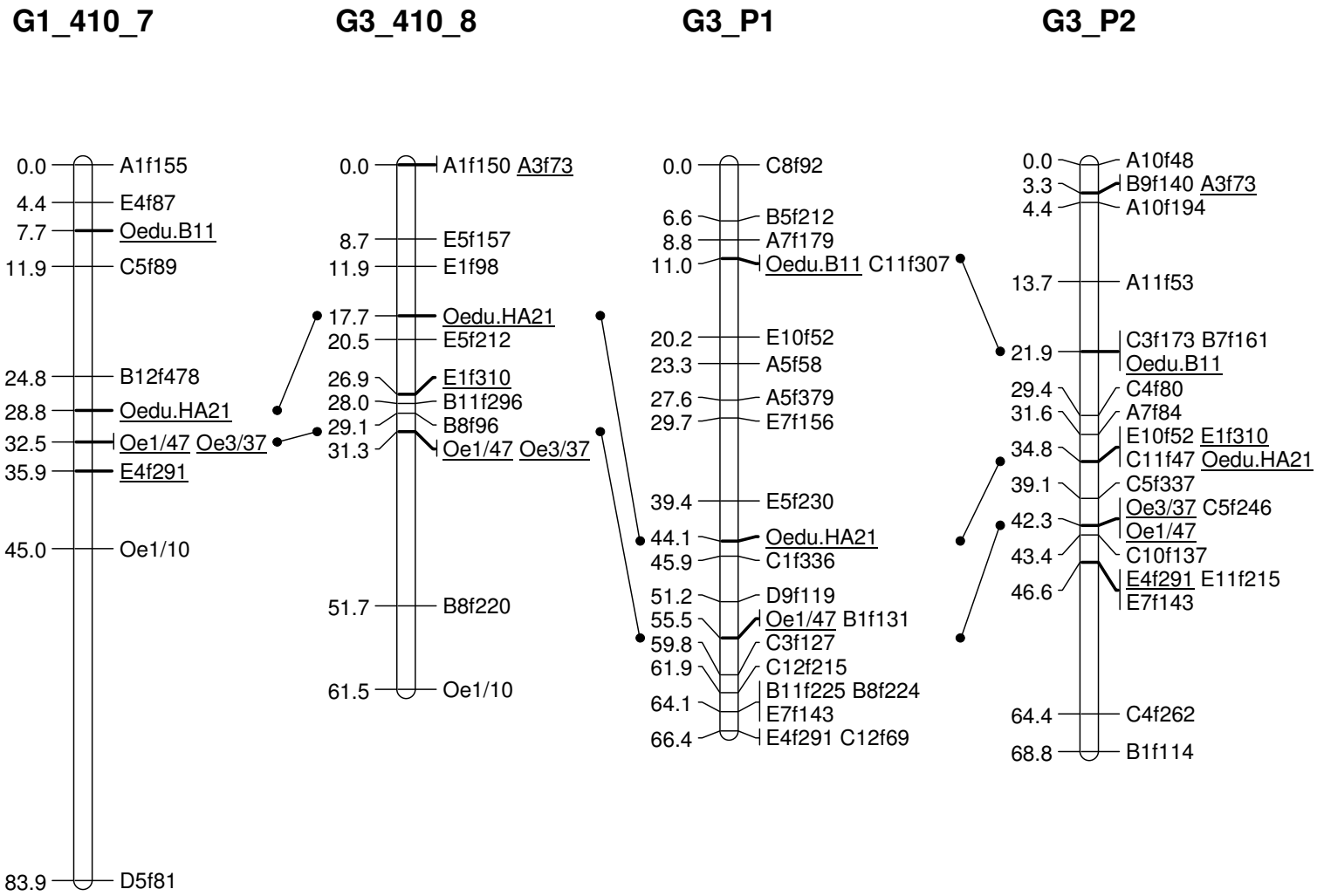


Figure 93. Homology group 1. Comparison of markers orders between the parental maps obtained in two different mapping families (OE.F2.04.63: maps P1 and P2; OE.F2.05.04: maps 410_7 and 410_8). Common markers are underlined.

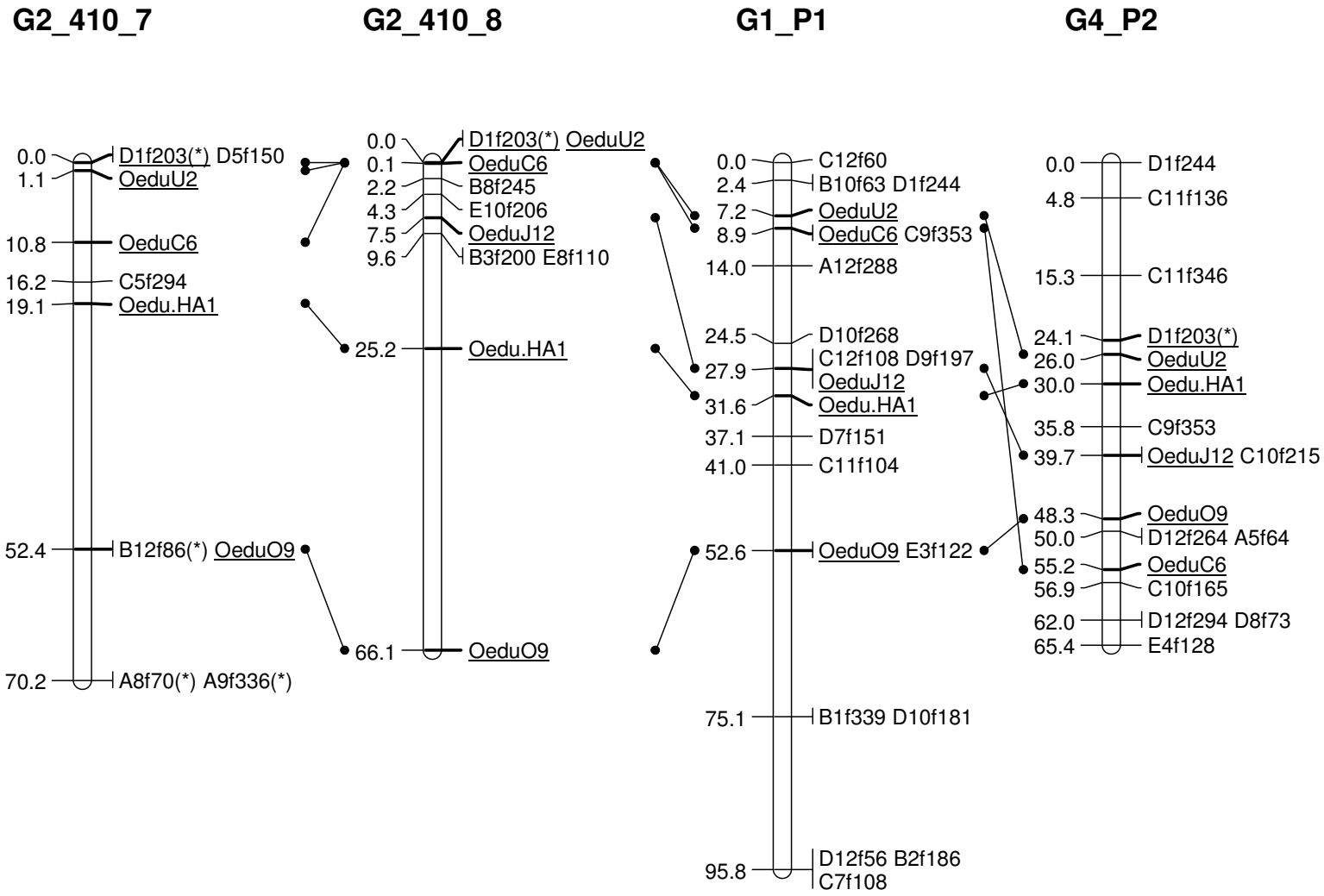


Figure 94. Homology group 2. Comparison of markers orders between the parental maps obtained in two different mapping families (OE.F2.04.63: maps P1 and P2; OE.F2.05.04: maps 410_7 and 410_8). Common markers are underlined.

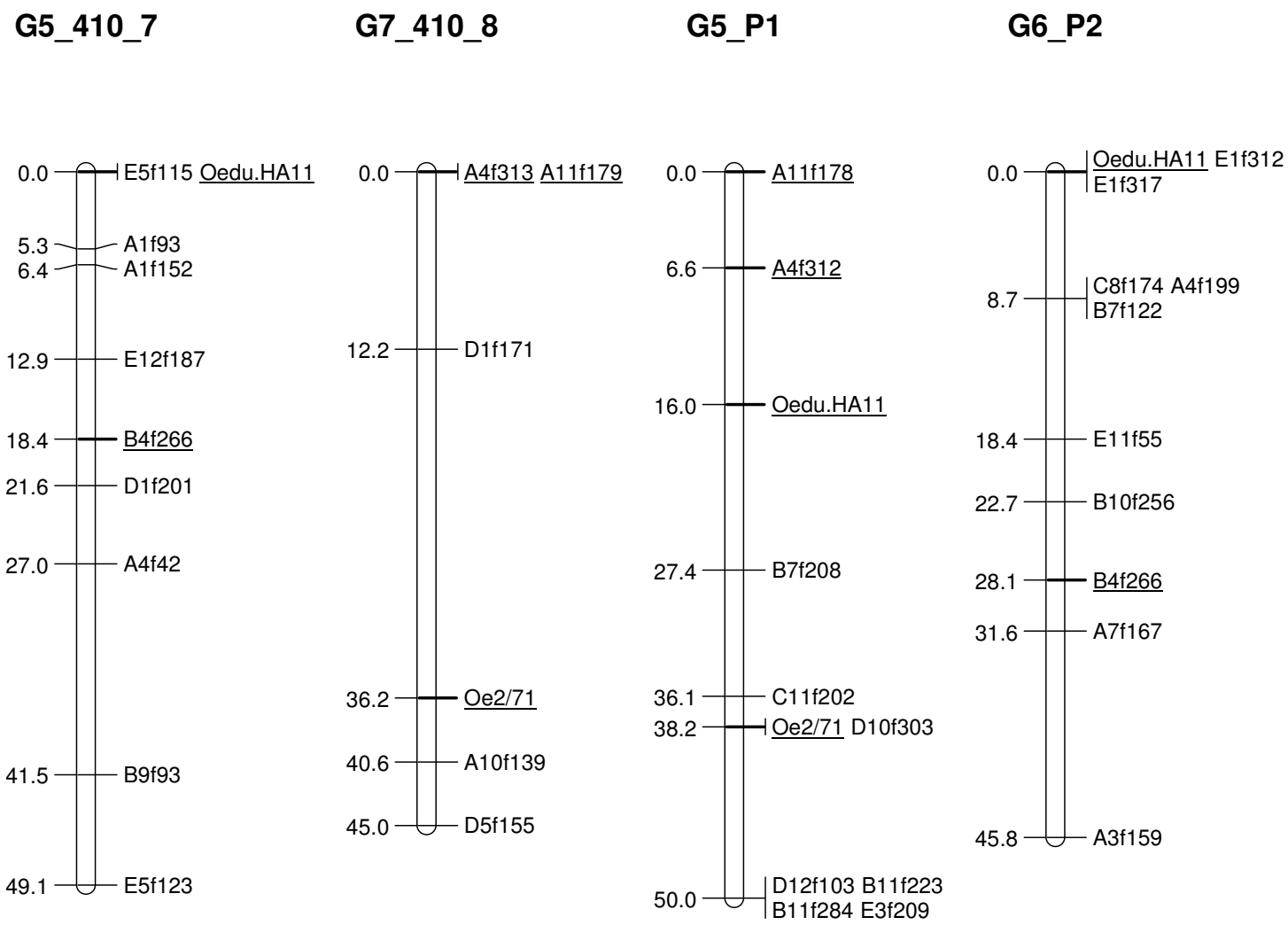


Figure 95. Homology group 3. Comparison of markers orders between the parental maps obtained in two different mapping families (OE.F2.04.63: maps P1 and P2; OE.F2.05.04: maps 410_7 and 410_8). Common markers are underlined.

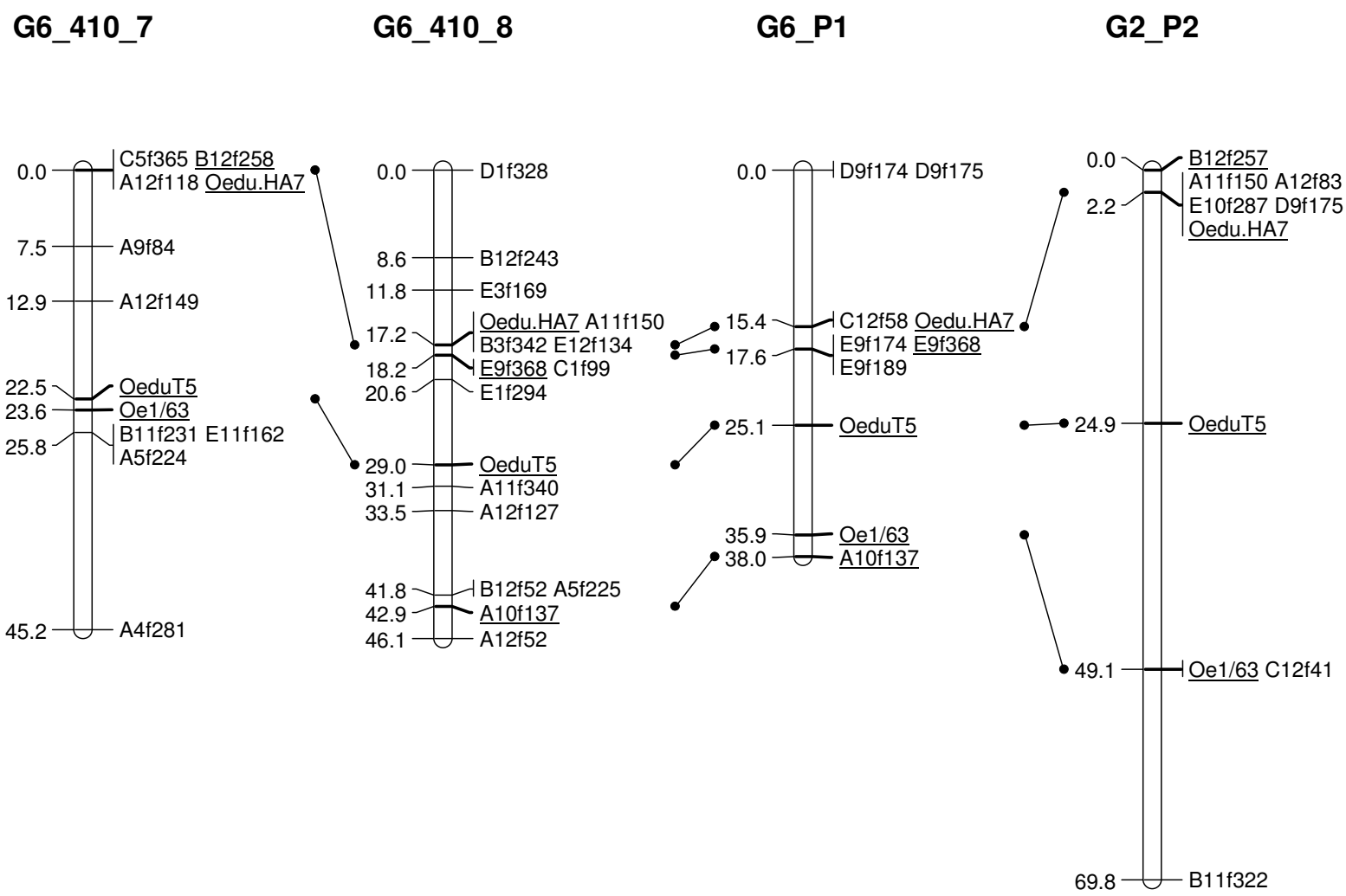


Figure 96. Homology group 4. Comparison of markers orders between the parental maps obtained in two different mapping families (OE.F2.04.63: maps P1 and P2; OE.F2.05.04: maps 410_7 and 410_8). Common markers are underlined.

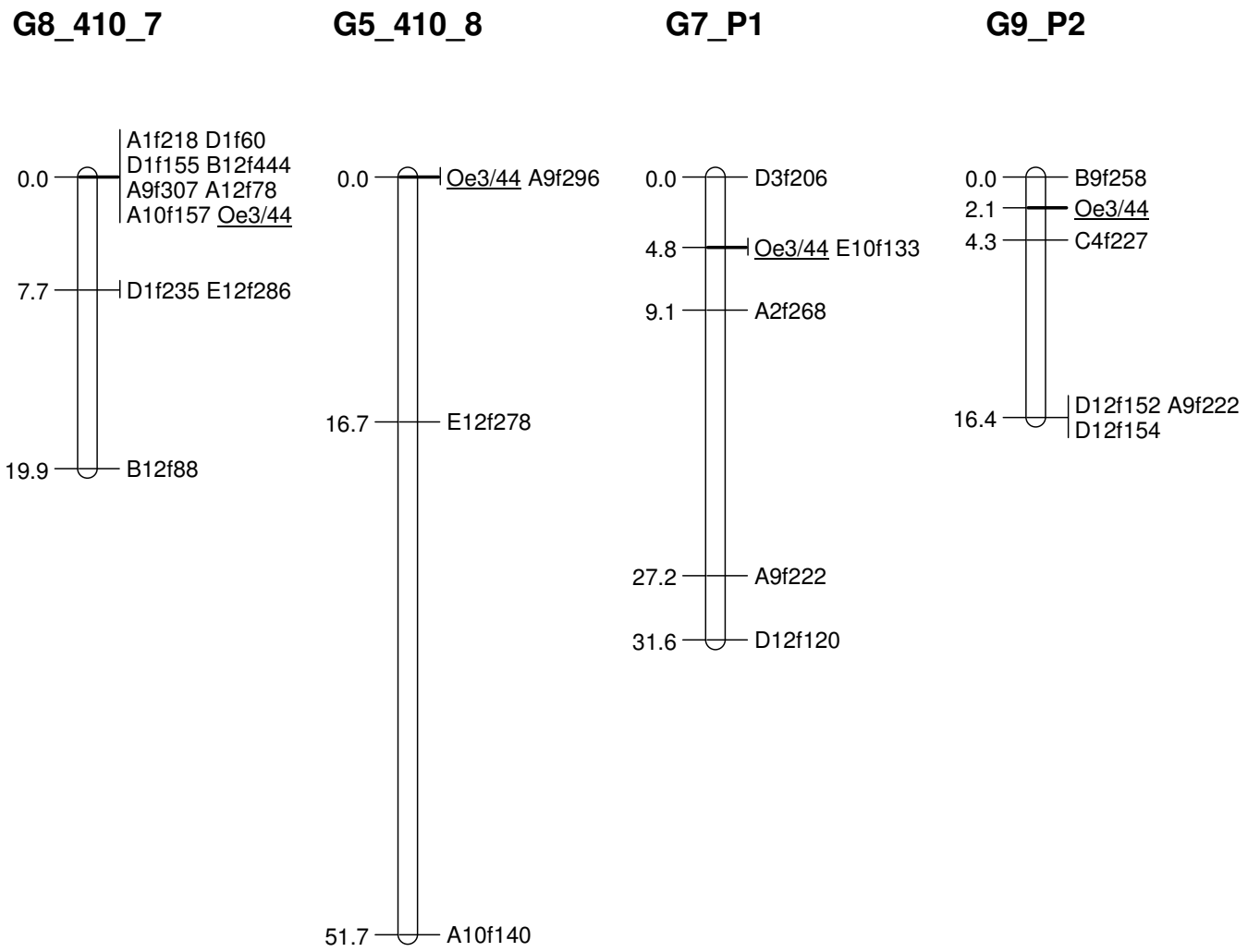


Figure 97. Homology group 5. Comparison of markers orders between the parental maps obtained in two different mapping families (OE.F2.04.63: maps P1 and P2; OE.F2.05.04: maps 410_7 and 410_8). Common markers are underlined.

IV- *BONAMIA* CHALLENGE EXPERIMENT AND QTL MAPPING

Two families were part of the challenge experiment: OE.F2.04.45 (self-fertilised family, 2-year old oysters) and OE.F2.05.04 (one-year old oysters). They were chosen on the basis of DNA polymorphism of the grand-parents and parents oysters, prior to performing genetic mapping in these two families. These families consisted of three-generation pedigree because this experimental design proved to be efficient for QTL mapping.

The *Bonamia* challenge experiment successfully demonstrated that the disease was transmitted from the wild oysters to the tested oysters, with the first mortalities in the tested oysters occurring after four months of cohabitation. Interestingly, higher mortalities and higher level of infection were observed in the one-year old family. This was inconsistent with the findings of Culloty and Mulcahy (1996) that 2 years is the critical age for the disease development. This unexpected result concerning the development of the disease in the two families studied was probably related to a different genetic background and to chance.

Despite the lack of power of our experimental design (family structure, small family size for the context of QTL mapping, dominance of molecular markers and low transferability, low number of progeny genotyped), interesting and promising results were obtained. Indeed, there was a good concordance in the results obtained with three different methodologies used: multi-stage testing strategy (Moen et al., 2004a), genetic mapping and QTL mapping with QTL express software. Several potential QTLs of resistance and susceptibility to bonamiasis were found. The results, even if preliminary, represent a first step towards MAS in the flat oyster.

V- FINAL CONSIDERATIONS

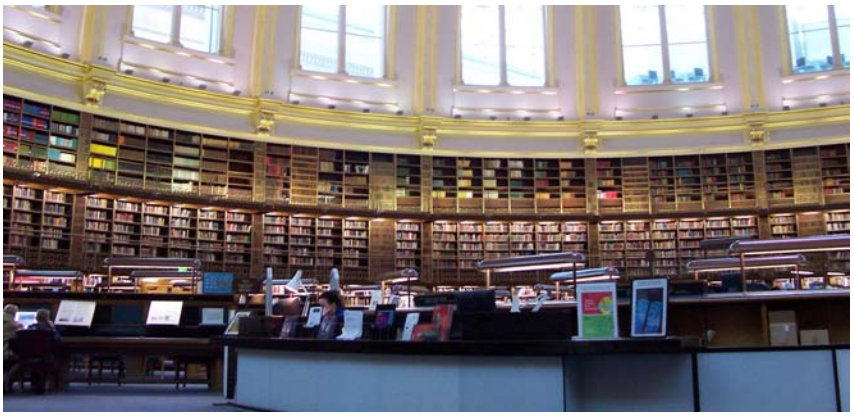
During this PhD project, only a few microsatellites were used and efforts were put into AFLP markers to obtain a sufficient number of genetic markers in the context of genetic and QTL mapping. However, if I had to start again, I would invest time developing microsatellites instead of using AFLPs. Indeed, even if AFLPs represent valuable genetic resources, their transferability from one family to another is limited and their dominance feature complicates

linkage analysis as well as the model to be used for QTL mapping. Therefore, microsatellites are worth the effort.

With hindsight, parentage analyses should have been performed before hand, in order to analyse the same family for genetic and QTL mapping. This is a very important point. Concentrating our efforts on a single family would have allowed us to increase the number of progeny genotyped. Erickson et al. (2004) stated that 300 progeny are appropriate for 10-15 cM marker intervals in most experimental designs to estimate the true distribution of QTL effects and that it is better to maximise sample size at the expense of marker density. However, despite the limited number of offspring genotyped in the QTL mapping experiment, we tried to increase the power of our experimental design by performing a kind of selective genotyping (we scored 46 progeny in two extreme phenotypic classes: dead oysters heavily infected with the parasite, surviving oysters not infected). This approach proved to be efficient because several QTLs of resistance/susceptibility to bonamiasis were found.

Finally, it is clear that there is a value in identifying the QTLs of resistance/susceptibility to bonamiasis in the context of marker-assisted selection. In the case of a tight association of a marker or a few markers to the disease resistance, the selection of progenitors based on marker genotypes would be a very valuable approach that would assist the breeder and speed up the selection process. Of course, MAS would be highly efficient if only a few QTLs explained a high proportion of the resistance to the disease. On the contrary, if many QTLs are controlling disease resistance, we can see that the application of MAS would be less valuable. Before achieving this goal, the road is still very long. Confirmatory experiments should be performed in other families. But this area of research represents a very promising approach towards the selection of resistant oysters and could lead to the revival of the flat oyster aquaculture industry in Europe.

REFERENCES



Agresti, J.J., Seki, S., Cnaani, A., Poompuang, S., Hallerman, E.M., Umiel, N., Hulata, G., Gall, G.A.E., May, B., 2000. Breeding new strains of tilapia: development of an artificial center of origin and linkage map based on AFLP and microsatellite loci. Aquaculture 185, 43-56.

Ajmone-Marsan, P., Valentini, A., Cassandro, M., Vecchiotti-Antaldi, G., Bertoni, G., Kuiper, M., 1997. AFLP markers for DNA fingerprinting in cattle. Animal Genetics 28, 418-426.

Alvarez, G., Zapata, C., Amaro, R., Guerr, A., 1989. Multilocus heterozygosity at protein loci and fitness in the European oyster, *Ostrea edulis* L. Heredity 63, 359-372.

Armour, J.A., Neumann, R., Gobert, S., Jeffreys, A.J., 1994. Isolation of human simple repeat loci by hybridization selection. Human Molecular Genetics 3, 599-565.

Balouet, G., Poder, J., Cahour, A., 1983. Haemocytic parasitosis: morphology and pathology of lesions in the French flat oyster, Ostrea edulis L. Aquaculture 34, 1-14

Baranski, M., Loughnan, S., Austin, C.M., Robinson, N., 2006. A microsatellite linkage map of the blacklip abalone, *Haliotis rubra*. Animal Genetics 37, 563-570.

Barreneche, T., Bodenes, C., Lexer, C., Trontin, J.-F., Fluch, S., Streiff, R., Plomion, C., Roussel, G., Steinkellner, H., Burg, K., Favre, J.-M., Glössl, J., Kremer, A., 1998. A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. Theoretical and Applied Genetics 97, 1090-1103.

Basten, C., Weir, B.S., Zeng, Z.-B., 1997. QTL Cartographer: a reference manual and tutorial for QTL mapping. Raleigh, N.C.: Department of Statistics, North Carolina State University.

Baud, J.-P., Gérard, A., Naciri-Graven, Y., 1997. Comparative growth and mortality of *Bonamia ostreae*-resistant and wild flat oysters, *Ostrea edulis*, in an intensive system. I. First year of experiment. Marine Biology 130, 71-79.

Bayne, B.L., 1965. Growth and the delay of metamorphosis of the larvae of *Mytilus edulis* (L.). Ophelia 2, 1-47.

Bayne, B.L., 1971. Some morphological changes that occur at the metamorphosis of the larvae of *Mytilus edulis*. In: (Editor), D.J.C. (Ed.), Proceedings of the 4th European Marine Biology Symposium. Cambridge University Press, London, Bangor, U.K., 1969, pp. 259-280.

Beaumont, A.R., 1991. Genetic studies of laboratory reared mussels, *Mytilus edulis*: heterozygote deficiencies, heterozygosity and growth. Biological Journal of the Linnean Society 44, 273-285.

Beaumont, A.R., 1994a. Linkage studies in *Mytilus edulis*, the mussel. Heredity 72, 557-562.

Beaumont, A.R., 1994b. The application and relevance of genetics in aquaculture. In: Beaumont, A.R. (Ed.), Genetics and evolution of aquatic organisms. Chapman and Hall, London.

Beaumont, A.R., Fairbrother, J.E., 1991. Ploidy manipulation in molluscan shellfish: a review. Journal of Shellfish Research 10, 1-18.

Beaumont, A.R., Seed, R., Garcia-Martinez, P., 1989. Electrophoretic and morphometric criteria for the identification of the mussels *Mytilus edulis* and *M. galloprovincialis*. In: Tyler, J.R.a.P.A. (Ed.), Proceedings of the 23rd European Marine Biology Symposium, Swansea, U.K. Olsen and Olsen, Fredensborg, Denmark, pp. 251-258.

Beaumont, A.R., Trebano Garcia, M., Hönig, S., Low, P., 2006. Genetics of Scottish populations of the native oyster, *Ostrea edulis*: gene flow, human intervention and conservation. Aquatic Living Resources 19, 389-402.

Bedier, E., Cochennec-Laureau, N., Langlade, A., Kopp, J., Goyard, E., Gerard, A., 2001. Recovery of the European flat oyster *Ostrea edulis* (L.): new development. Proceedings of the European Aquaculture Society annual meeting 2001.

Behura, S.K., Valicente, F.H., Rider, S.D., Jr., Shun-Chen, M., Jackson, S., Stuart, J.J., 2004. A physically anchored genetic map and linkage to avirulence reveals recombination suppression over the proximal region of Hessian fly chromosome A2. Genetics 167, 343-355.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F., 1996-2001. GENETIX 4.02, logiciel sous Windows TM pour la génétique des populations. , Montpellier, France: Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II.

Benzie, J.A.H., Kenway, M., Trott, L., 1997. Estimates for the heritability of size in juvenile *Penaeus monodon* prawns from half-sib matings. Aquaculture 152, 49-53.

Berthe, F.C.J., Le Roux, F., Adlard, R.D., Figueras, A., 2004. Marteiliosis in molluscs: A review. Aquatic Living Resources 17, 433-448.

Bierne, N., Launey, S., Naciri-Graven, Y., Bonhomme, F., 1998. Early effect of inbreeding as revealed by microsatellite analyses on *Ostrea edulis* larvae. Genetics 148, 1893-1906.

Boudry, P., Collet, B., Cornette, F., Hervouet, V., Bonhomme, F., 2002. High variance in reproductive success of the Pacific oyster (*Crassostrea gigas*, Thunberg) revealed by microsatellite-based parentage analysis of multifactorial crosses. Aquaculture 204, 283-296.

Bourget, E., 1983. Seasonal variations of cold tolerance in intertidal molluscs and their relation to environmental conditions in the St. Lawrence Estuary. Canadian Journal of Zoology 61, 1193-1201.

Bower, S.M., 1992. Diseases and parasites of mussels. In: Gosling, E.M. (Ed.), The mussel *Mytilus*: ecology, physiology, genetics and culture. Elsevier Science Publ., Amsterdam, pp. 543-563.

Bower, S.M., McGladdery, S.E., 1996. Synopsis of Infectious Diseases and Parasites of Commercially Exploited Shellfish. SeaLane Diseases of Shellfish. Available from: <http://www-sci.pac.dfo-mpo.gc.ca/sealane/aquac/pages/toc.htm>.

Brehelin, M., Bonami, J.R., Cousserans, F., Vivares, C.P., 1982. Existence de formes plasmodiales vraies chez *Bonamia ostreae* parasite de l'huitre plate *Ostrea edulis*. Compte Rendus Hebdomadaires des Seances de l'Academie des Sciences, Serie D, Sciences Naturelles 295, 45-48.

Brichette, I., Reyero, M.I., Garcia, C., 2001. A genetic analysis of intraspecific competition for growth in mussel cultures. Aquaculture 192, 155-169.

Brookfield, J.F., 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5, 453-455.

Brown, J., Sundaresan, V., 1991. A recombination hotspot in the maize A1 intragenic region. Theoretical and Applied Genetics 81, 185-188.

Burr, B., Burr, F.A., Thompson, K.H., Albertson, M.C., Stuber, C.W., 1988. Gene mapping with recombinant inbreds in maize. Genetics 118, 519-526.

Caceres-Martinez, J., Robledo, J.A.F., Figueras, A., 1995. Presence of *Bonamia* and its relation to age, growth rates and gonadal development of the flat oyster, *Ostrea edulis*, in the Ria de Vigo, Galicia (NW Spain). Aquaculture 130, 15-23.

Callen, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C., Sutherland, G.R., 1993. Incidence and origin of "null" alleles in the (AC)n microsatellite markers. American Journal of Human Genetics 52, 922-927.

Campbell, D., Duchesne, P., Bernatchez, L., 2003. AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. Molecular Ecology 12, 1979-1991.

Carmelli, D., Karlin, S., Williams, R.R., 1979. A class of indices to assess major-gene versus polygenic inheritance of distributive variables. Progress in Clinical and Biological Research 32, 259-270.

Carnegie, R.B., Barber, B.J., Distel, D.L., 2003. Detection of the oyster parasite *Bonamia ostreae* by fluorescent in situ hybridization. Diseases of Aquatic Organisms 55, 247-252.

Carnegie, R.B., Barber, B.J., Culloty, S.C., Figueras, A.J., Distel, D.L., 2000. Development of a PCR assay for detection of the oyster pathogen *Bonamia ostreae* and support for its inclusion in the Haplosporidia. Diseases of Aquatic Organisms 42, 199-206.

Castiglioni, P., Ajmone-Marsan, P., van Wijk, R., Motto, M., 1999. AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theoretical and Applied Genetics 99, 425-431.

Ceccherelli, V.U., Rossi, R., 1984. Settlement, growth and production of the mussel *Mytillus galloprovincialis*. Marine Ecology Progress Series 16, 173-184.

Chagné, D., Lalanne, C., Madur, D., Kumar, S., Frigério, J.-M., Krier, C., Decroocq, S., Savouré, A., Bou-Dagher-Kharrat, M., Bertocchi, E., Brach, J., Plomion, C., 2002. A high density genetic map of maritime pine based on AFLPs. Annals of Forest Science 59, 627-636.

Chakraborty, R., De Andrade, M., Daiger, S.P., Budowle, B., 1992. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. Annals of Human Genetics 56 (Pt 1), 45-57.

Chakraborty, R., Kimmel, M., Stivers, D.N., Davison, L.J., Deka, R., 1997. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proceedings of the National Academy of Sciences of the USA 94, 1041-1046.

Chakravarti, A., Lasher, L.K., Reefer, J.E., 1991. A maximum likelihood method for estimating genome length using genetic linkage data. Genetics 128, 175-182.

Chambers, G.K., MacAvoy, E.S., 2000. Microsatellites: consensus and controversy. Comparative Biochemistry and Physiology part B 126, 455-476.

Cheng, Z., Presting, G.G., Buell, C.R., Wing, R.A., Jiang, J., 2001. High-resolution pachytene chromosome mapping of bacterial artificial chromosomes anchored by genetic markers reveals the centromere location and the distribution of genetic recombination along chromosome 10 of rice. Genetics 157, 1749-1757.

Churchill, G.A., Doerge, R.W., 1994. Empirical threshold values for quantitative trait mapping. Genetics 138, 963-971.

Cnaani, A., Zilberman, N., Tinman, S., Hulata, G., Ron, M., 2004. Genome-scan analysis for quantitative trait loci in an F2 tilapia hybrid. Molecular Genetics and Genomics 272, 162-172.

Cnaani, A., Hallerman, E.M., Ron, M., Weller, J.I., Indelman, M., Kashi, Y., Gall, G.A.E., Hulata, G., 2003. Detection of a chromosomal region with two quantitative trait loci, affecting cold tolerance and fish size, in an F2 tilapia hybrid. Aquaculture 223, 117-128.

Cochennec-Laureau, N., Auffret, M., Renault, T., Langlade, A., 2003. Changes in circulating and tissue-infiltrating hemocyte parameters of European flat oysters, *Ostrea edulis*, naturally infected with *Bonamia ostreae*. Journal of Invertebrate Pathology 83, 23-30.

Cochennec, N., Le Roux, F., Berthe, F., Gerard, A., 2000. Detection of *Bonamia ostreae* based on small subunit ribosomal probe. Journal of Invertebrate Pathology 76, 26-32.

Coimbra, M.R.M., Kobayashi, K., Koretsugu, S., Hasegawa, O., Ohara, E., Ozaki, A., Sakamoto, T., Naruse, K., Okamoto, N., 2003. A genetic linkage map of the Japanese flounder, *Paralichthys olivaceus*. Aquaculture 220, 203-218.

Collet, B., Boudry, P., Thebault, A., Heurtebise, S., Morand, B., Gerard, A., 1999. Relationship between pre- and post-metamorphic growth in the Pacific oyster *Crassostrea gigas* (Thunberg). Aquaculture 175, 215-226.

Collevatti, R.G., Grattapaglia, D., Hay, J.D., 2001. Population genetic structure of the endangered tropical tree species *Caryocar brasiliense*, based on variability at microsatellite loci. Molecular Ecology 10, 349-356.

Coltman, D.W., Bowen, W.D., Wright, J.M., 1998. Male mating success in an aquatically mating pinniped, the harbour seal (*Phoca vitulina*), assessed by microsatellite DNA markers. Molecular Ecology 7, 627-638.

Comps, M., Tige, G., Grizel, H., 1980. Etude ultrastructurale d'un Protiste parasite de l'huitre *Ostrea edulis* L. Compte Rendus Hebdomadaires des Seances de l'Academie des Sciences de Paris, t. 290, Serie D, 383-385.

Corbeil, S., Arzul, I., Diggles, B., Heasman, M., Chollet, B., Berthe, F.C.J., Crane, M.S.J., 2006. Development of a TaqMan PCR assay for the detection of *Bonamia* species. Diseases of Aquatic Organisms 71, 75-80.

Cornet, M., 1993. A short term culture for chromosome preparation from somatic tissues of adult mussels. Experientia 49, 87-90.

Crenshaw, J.W., Heffernan, P.B., Walker, R.L., 1991. Heritability of growth rate in the southern bay scallop, *Argopecten irradians concentricus* (Say, 1822). Journal of Shellfish Research 10, 55-63.

Crenshaw, J.W., Heffernan, P.B., Walker, R.L., 1996. Effect of growout density on heritability of growth rate in the Northern Quahog, *Mercenaria mercenaria* (Linnaeus, 1758). Journal of Shellfish Research 15, 341-344.

Crnokrak, P., Barrett, S.C., 2002. Perspective: purging the genetic load: a review of the experimental evidence. Evolution 56, 2347-2358.

Culloty, S.C., Mulcahy, M.F., 1996. Season-, age-, and sex-related variation in the prevalence of bonamiasis in flat oysters (*Ostrea edulis* L.) on the south coast of Ireland. Aquaculture 144, 53-63.

Culloty, S.C., Cronin, M.A., Mulcahy, M.F., 2003. Possible limitations of diagnostic methods recommended for the detection of the protistan, *Bonamia ostreae* in the European flat oyster, *Ostrea edulis*. Bulletin of the European Association of Fish Pathologists 23, 67-71.

Culloty, S.C., Novoa, B., Pernas, M., Longshaw, M., Mulcahy, M.F., Feist, S.W., Figueras, A., 1999. Susceptibility of a number of bivalve species to the protozoan parasite *Bonamia ostreae* and their ability to act as a vector for this parasite. Diseases of Aquatic Organisms 37, 73-80.

D'Amato, M.E., Lunt, D.H., Carvalho, G.R., 1999. Microsatellite markers for the hake *Macruronus magellanicus* amplify other gadoid fish. Molecular Ecology 8, 1086-1088.

da Silva, P.M., Villalba, A., 2004. Comparison of light microscopic techniques for the diagnosis of the infection of the European flat oyster *Ostrea edulis* by the protozoan *Bonamia ostreae*. Journal of Invertebrate Pathology 85, 97-104.

Daguin, C., Borsa, P., 1999. Genetic characterisation of *Mytilus galloprovincialis* Lmk. in North West Africa using nuclear DNA markers. Journal of Experimental Marine Biology and Ecology 235, 55-65.

Dallas, J.F., 1992. Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. Mammalian Genome 3, 452-456.

Danzmann, R.G., Jackson, T.R., M. Ferguson, M., 1999. Epistasis in allelic expression at upper temperature tolerance QTL in rainbow trout. Aquaculture 173, 45-58.

Darvasi, A., Soller, M., 1992. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theoretical and Applied Genetics 85, 353-359.

Darvasi, A., Weinreb, A., Minke, V., Weller, J.I., Soller, M., 1993. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics 134, 943-951.

David, P., 1998. Heterozygosity-fitness correlations: new perspectives on old problems. Heredity 80, 531–537.

David, P., Delay, B., Berthou, P., Jarne, P., 1995. Alternative models for allozyme-associated heterosis in the marine bivalve *Spisula ovalis*. Genetics 139, 1719-1726.

Davis, C.V., 2000. Estimation of narrow-sense heritability for larval and juvenile growth traits in selected and unselected sub-lines of eastern oysters, *Crassostrea virginica*. Journal of Shellfish Research 19, 613.

de Gortari, M.J., Freking, B.A., Cuthbertson, R.P., Kappes, S.M., Keele, J.W., Stone, R.T., Leymaster, K.A., Dodds, K.G., Crawford, A.M., Beattie, C.W., 1998. A second-generation linkage map of the sheep genome. Mammalian Genome 9, 204-209.

Dégremont, L., Ernande, B., Bédier, E., Boudry, P., 2007. Summer mortality of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). I. Estimation of genetic parameters for survival and growth. Aquaculture 262, 41-53.

Diaz-Almela, E., Boudry, P., Launey, S., Bonhomme, F., Lapegue, S., 2004. Reduced female gene flow in the European flat oyster *Ostrea edulis*. Journal of Heredity 95, 510-516.

Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., Weissenbach, J., 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380, 152-154.

Dietrich, W.F., Miller, J., Steen, R., Merchant, M.A., Damron-Boles, D., Husain, Z., Dredge, R., Daly, M.J., Ingalls, K.A., O'Connor, T.J., 1996. A comprehensive genetic map of the mouse genome. Nature 380, 149-152.

Di Rienzo, A., Peterson, A.C., Garza, J.C., Valdes, A.M., Slatkin, M., Freimer, N.B., 1994. Mutational processes of simple-sequence repeat loci in human populations. Proceedings of the National Academy of Sciences of the USA 91, 3166-3170.

Donini, P., Elias, M.L., Bougourd, S.M., Koebner, R.M.D., 1997. AFLP fingerprinting reveals pattern differences between template DNA extracted from different plant organs. Genome 40, 521-526.

Edwards, K.J., Barker, J.H., Daly, A., Jones, C., Karp, A., 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. Biotechniques 20, 758-760.

Ellegren, H., 2004. Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics 5, 435-445.

Ellstrand, N.C., 1984. Multiple paternity within the fruits of the wild radish, *Raphanus sativus*. The American Naturalist 123, 819-828.

Elston, R.A., Farley, C.A., Kent, M.L., 1986. Occurrence and significance of bonamiasis in European flat oysters *Ostrea edulis* in North America. Diseases of Aquatic Organisms 2, 49-54.

Elston, R.A., Kent, M.L., Wilkinson, M.T., 1987. Resistance of *Ostrea edulis* to *Bonamia ostreae* infection. Aquaculture 64, 237-242.

Elvingson, P., Johansson, K., 1993. Genetic and environmental components of variation in body traits of rainbow trout (*Oncorhynchus mykiss*) in relation to age. Aquaculture 118, 191-204.

Erickson, D.L., Fenster, C.B., Stenoien, H.K., Price, D., 2004. Quantitative trait locus analyses and the study of evolutionary process. Molecular Ecology 13, 2505-2522.

Escaravage, N., Questiau, S., Pornon, A., Doche, B., Taberlet, P., 1998. Clonal diversity in a *Rhododendron ferrugineum* L. (Ericaceae) population inferred from AFLP markers. Molecular Ecology 7, 975-982.

Estoup, A., Garnery, L., Solignac, M., Cornuet, J.M., 1995a. Microsatellite variation in honey bee (*Apis mellifera* L.) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. Genetics 140, 679-695.

Estoup, A., Tailliez, C., Cornuet, J.M., Solignac, M., 1995b. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). Molecular Biology and Evolution 12, 1074-1084.

Eyster, L.S., Pechenik, J.A., 1987. Attachment of *Mytilus edulis* L. larvae on algal and byssal filaments is enhanced by water agitation. Journal of Experimental Marine Biology and Ecology 114, 99-110.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to quantitative genetics. Longman, Harlow, England, 464 pp.

Faris, J.D., Haen, K.M., Gill, B.S., 2000. Saturation mapping of a gene-rich recombination hot spot region in wheat. Genetics 154, 823-835.

Figueras, A.J., 1991. *Bonamia* status and its effects in cultured flat oysters in the Ria de Vigo, Galicia (N.W. Spain). Aquaculture 93, 225-233.

Foltz, D.W., 1986. Null alleles as a possible cause of heterozygote deficiencies in the oyster *Crassostrea virginica* and other bivalves. Evolution 40, 869-870.

Friedman, C.S., Perkins, F.O., 1994. Range extension of *Bonamia ostreae* to Maine, U.S.A. Journal of Invertebrate Pathology 64, 179-181.

Friedman, C.S., McDowell, T., Groff, J.M., Hollibaugh, J.T., Manzer, D., Hedrick, R.P., 1989. Presence of *Bonamia ostreae* among populations of the European flat oyster, *Ostrea edulis* L., in California, USA. Journal of Shellfish Research 8, 133-137.

Fu, Y.B., 1999. Patterns of the purging of deleterious genes with synergistic interactions in different breeding schemes Theoretical and Applied Genetics 98, 337-346.

Garza, J.C., Slatkin, M., Freimer, N.B., 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Molecular Biology and Evolution 12, 594-603.

Gates, M.A., Kim, L., Egan, E.S., Cardozo, T., Sirotkin, H.I., Dougan, S.T., Lashkari, D., Abagyan, R., Schier, A.F., Talbot, W.S., 1999. A genetic linkage map for zebrafish: comparative analysis and localization of genes and expressed sequences. Genome Research 9, 334-347.

Gerber, S., Mariette, S., Streiff, R., Bodenes, C., Kremer, A., 2000. Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. Molecular Ecology 9, 1037-1048.

Gjedrem, T., 1983. Genetic variation in quantitative traits and selective breeding in fish and shellfish. Aquaculture 33, 51-72.

Gjedrem, T., 2000. Genetic improvement of cold-water fish species. Aquaculture Research 31, 25-33.

Gjerde, B., Gjedrem, T., 1984. Estimates of phenotypic and genetic parameters for carcass traits in Atlantic salmon and rainbow trout. Aquaculture 36, 97-110.

Gjerde, B., Schaeffer, L.R., 1989. Body traits in rainbow trout: II. Estimates of heritabilities and of phenotypic and genetic correlations. Aquaculture 80, 25-44.

Gonzalez-Tizon, A.M., Martinez-Lage, A., Rego, I., Ausio, J., Mendez, J., 2000. DNA content, karyotypes, and chromosomal location of 18S-5.8S-28S ribosomal loci in some species of bivalve molluscs from the Pacific Canadian coast. Genome 43, 1065-1072.

Gosling, E.M., 1992. The mussel *Mytilus*: ecology, physiology, genetics, and culture. Elsevier, Amsterdam, 589 pp.

Goulletquer, P., Heral, M., 1997. Marine molluscan production trends in France: from fisheries to aquaculture. U.S. Dep Commer., NMFS 129, 137-164.

Grattapaglia, D., Sederoff, R., 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137, 1121-1137.

Green, S., Falls, K., Crooks, S., 1990. Documentation for CRI-MAP v2.4. Washington University School of Medicine, St. Louis.

Grewe, P.M., Krueger, C.C., Aquadro, C.F., Bermingham, E., Kincaid, H.L., May, B., 1993. Mitochondrial DNA variation among lake trout (*Salvelinus namaycush*) strains stocked into Lake Ontario. Canadian Journal of Fisheries and Aquatic Sciences 50, 2397-2403.

Grizel, H., 1985. Etude des récentes épizooties de l'huître plate *Ostrea edulis* Linné et de leur impact sur l'ostréiculture bretonne. Thèse Doctoral d'Etat en Sciences Naturelles, Université de Montpellier, 145pp.

Groenen, M.A., Cheng, H.H., Bumstead, N., Benkel, B.F., Briles, W.E., Burke, T., Burt, D.W., Crittenden, L.B., Dodgson, J., Hillel, J., Lamont, S., de Leon, A.P., Soller, M., Takahashi, H., Vignal, A., 2000. A consensus linkage map of the chicken genome. Genome Research 10, 137-147.

Guo, X., Gaffney, P.M., 1993. Artificial gynogenesis in the Pacific oyster, *Crassostrea gigas*: II. Allozyme inheritance and early growth. The Journal of Heredity 84, 311-315.

Guo, X., Hershberger, W.K., Cooper, K., Chew, K.K., 1993. Artificial gynogenesis with ultraviolet light-irradiated sperm in the Pacific oyster, *Crassostrea gigas*. I. Induction and survival. Aquaculture 113, 201-214.

Hadley, N.H., Dillon, J.R.T., Manzi, J.J., 1991. Realized heritability of growth rate in the hard clam Mercenaria mercenaria. Aquaculture 93, 109-119.

Haldane, J.B.S., 1919. The combination of linkage values, and the calculation of distances between the loci of linked factors. Journal of genetics 8, 299-309.

Haley, C.S., Knott, S.A., 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69, 315-324.

Hamada, H., Kakunaga, T., 1982. Potential Z-DNA forming sequences are highly dispersed in the human genome. Nature 298, 396-398.

Hancock, J.M., 1999. Microsatellites and other simple sequences: genome context and mutational mechanisms. In: Goldstein, D., Schlotterer, C. (Ed.), Microsatellites. Evolution and applications. Oxford University Press, New York, pp. 1-9.

Hardy, O.J., 2003. Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. Molecular Ecology 12, 1577-1588.

Harger, J.R.E., Landenberger, D.E., 1971. The effects of storms as a density dependent mortality factor on populations of sea mussels. The Veliger 14, 195-210.

Hearne, C.M., Ghosh, S., Todd, J.A., 1992. Microsatellites for linkage analysis of genetic traits. Trends in Genetics 8, 288-294.

Hedgecock, D., Cooper, K., Hershberger, B., Guo, X., 1991. Body size at harvest and sex ratio of pacific oysters *Crassostrea gigas* from controlled crosses. Journal of Shellfish Research 10, 237.

Hedgecock, D., Gaffney, P.M., Goulletquer, P., Guo, X., Reece, K., Warr, G.W., 2005. The case for sequencing the Pacific oyster genome. Journal of Shellfish Research 24, 429-441.

Hedgecock, D., Launey, S., Pudovkin, A.I., Naciri-Graven, Y., Lapègue, S., Bonhomme, F., 2007. Small effective number of parents (Nb) inferred for a naturally settled cohort of juvenile European flat oysters *Ostrea edulis*. In press in Marine Biology.

Hedrick, P.W., 2000. Genetics of populations- Second edition. Jones and Bartlett Publishers, 553 pp.

Henryon, M., Jokumsen, A., Berg, P., Lund, I., Pedersen, P.B., Olesen, N.J., Slierendrecht, W.J., 2002. Genetic variation for growth rate, feed conversion efficiency, and disease resistance exists within a farmed population of rainbow trout. Aquaculture 209, 59-76.

Herbinger, C., Doyle, R.W., Taggart, C.T., Lochmann, S.E., Brooker, A.L., Wright, J.M., Cook, D., 1997. Family relationships and effective population size in a natural cohort of Atlantic cod (*Gadus morhua*) larvae. Canadian Journal of Fisheries and Aquatic Sciences 54, 11-18.

Hervio, D., Bachere, E., Boulo, V., Cochennec, N., Vuillemin, V., Le Coguic, Y., Cailletaux, G., Mazurie, J., Mialhe, E., 1995. Establishment of an experimental infection protocol for the flat oyster, *Ostrea edulis*, with the intrahaemocytic protozoan parasite, *Bonamia ostreae*: application in the selection of parasite-resistant oysters. Aquaculture 132, 183-194.

Hetzel, D.J.S., Crocos, P.J., Davis, G.P., Moore, S.S., Preston, N.C., 2000. Response to selection and heritability for growth in the Kuruma prawn, *Penaeus japonicus*. Aquaculture 181, 215-223.

Hill, M., Witsenboer, H., Zabeau, M., Vos, P., Kesseli, R., Michelmore, R., 1996. PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. Theoretical and Applied Genetics 93, 1202-1210.

Hinegardner, R., 1974. Cellular DNA content of the Mollusca. Comparative Biochemistry and Physiology Part A 47, 447-460.

Howson, C.M., Picton, B.E., 1997. The species directory of the marine fauna and flora of the British Isles and surrounding seas. Ulster Museum and The Marine Conservation Society, Belfast and Ross-on-Wye, Ulster Museum publication, no. 276.

Hubert, S., Hedgecock, D., 2004. Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas*. Genetics 168, 351-362.

Insua, A., Labat, J.P., Thiriot-Quievreux, C., 1994. Comparative analysis of karyotypes and nucleolar organizer regions in different populations of *Mytilus trossulus*, *Mytilus edulis* and *Mytilus galloprovincialis*. Journal of Molluscan Studies 60, 359-360.

Jablonski, D.J., Lutz, R.A., 1980. Molluscan larval shell morphology: ecological and paleoecological applications. In: Lutz, D.C.R.a.R.A. (Ed.), Skeletal growth of aquatic organisms. Plenum Press, New York, pp. 323-377.

Jackson, T.R., Ferguson, M.M., Danzmann, R.G., Fishback, A.G., Ihssen, P.E., O'Connell, M., Crease, T.J., 1998. Identification of two QTL influencing upper temperature tolerance in three rainbow trout (*Oncorhynchus mykiss*) half-sib families. Heredity 80, 143-151.

Jansen, R.C., 1993. Interval mapping of multiple quantitative trait loci. Genetics 135, 205-211.

Jansen, R.C., 1996. A general Monte Carlo method for mapping multiple quantitative trait loci. Genetics 142, 305-311.

Jarayabhand, P., Thavornyutikarn, M., 1995. Realized heritability estimation on growth rate of oyster, *Saccostrea cucullata* Born, 1778. Aquaculture 138, 111-118.

Jarne, P., Lagoda, P.J.L., 1996. Microsatellites, from molecules to populations and back. Trends in Ecology & Evolution 11, 424-429.

Jaziri, H., 1985. Biogeographie et polymorphisme enzymatique chez l'huitre plate *Ostrea edulis* L. 1758. Memoire de DEA, USTL.

Jaziri, H., 1990. Variations genetiques et structuration biogeographique chez un bivalve marin: l'huitre plate *Ostrea edulis* L. (PhD thesis). University of Montpellier II, Montpellier, France.

Jeffreys, A.J., Wilson, V., Thein, S.L., 1985. Hypervariable 'minisatellite' regions in human DNA. Nature 314, 67-73.

Jenny, M.J., Ringwood, A.H., Lacy, E.R., Lewitus, A.J., Kempton, J.W., Gross, P.S., Warr, G.W., Chapman, R.W., 2002. Potential indicators of stress response identified by expressed sequence tag analysis of hemocytes and embryos from the American oyster, *Crassostrea virginica*. Marine Biotechnology 4, 81-93.

Jiang, C., Zeng, Z.B., 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140, 1111-1127.

Johannesson, K., Rodstrom, E.M., Aase, H., 1989. Low genetic variability in Scandinavian populations of *Ostrea edulis* L. - possible causes and implications. Journal of Experimental Marine Biology and Ecology 128, 177-190.

Johnson, S.L., Gates, M.A., Johnson, M., Talbot, W.S., Horne, S., Baik, K., Rude, S., Wong, J.R., Postlethwait, J.H., 1996. Centromere-linkage analysis and consolidation of the zebrafish genetic map. Genetics 142, 1277-1288.

Jonasson, J., 1993. Selection experiments in salmon ranching. I. Genetic and environmental sources of variation in survival and growth in freshwater. Aquaculture 109, 225-236.

Jonasson, J., Gjerde, B., Gjedrem, T., 1997. Genetic parameters for return rate and body weight of sea-ranched Atlantic salmon. Aquaculture 154, 219-231.

Jonasson, J., Stefansson, S.E., Gudnason, A., Steinarsson, A., 1999. Genetic variation for survival and shell length of cultured red abalone (*Haliotis rufescens*) in Iceland. Journal of Shellfish Research 18, 621-625.

Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., Van der Wiel, C., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., 1998. Reproducibility testing of AFLPs by a network of European laboratories. In: Karp, A., Isaac, P.G., Ingram, D.S. (Ed.), Molecular tools for screening biodiversity. Chapman and Hall, London, pp. 191-192.

Jones, R., Bates, J.A., Innes, D.J., Thompson, R.J., 1996. Quantitative genetic analysis of growth in larval scallops (*Placopecten magellanicus*). Marine Biology 124, 671-677.

Kandpal, R.P., Kandpal, G., Weissman, S.M., 1994. Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. Proceedings of the National Academy of Sciences of the USA 91, 88-92.

Kappes, S.M., Keele, J.W., Stone, R.T., McGraw, R.A., Sonstegard, T.S., Smith, T.P., Lopez-Corrales, N.L., Beattie, C.W., 1997. A second-generation linkage map of the bovine genome. Genome Research 7, 235-249.

Karagyozov, L., Kalcheva, I.D., Chapman, V.M., 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Research 21, 3911-3912.

Katagiri, T., Asakawa, S., Hirono, I., Aoki, T., Shimizu, N., 2000. Genomic bacterial artificial chromosome library of the Japanese flounder *Paralichthys olivaceus*. Marine Biotechnology 2, 571-576.

Katagiri, T., Asakawa, S., Minagawa, S., Shimizu, N., Hirono, I., Aoki, T., 2001. Construction and characterization of BAC libraries for three fish species; rainbow trout, carp and tilapia. Animal Genetics 32, 200-204.

Keane, B., Dittus, W.P., Melnick, D.J., 1997. Paternity assessment in wild groups of toque macaques *Macaca sinica* at Polonnaruwa, Sri Lanka using molecular markers. Molecular Ecology 6, 267-282.

Kearsey, M.J., Hyne, V., 1994. QTL analysis: a simple "marker-regression" approach. Theoretical and Applied Genetics 89, 698-702.

Kearsey, M.J., Pooni, H.S., 1998. The genetical analysis of quantitative traits. Stanley Thornes, Cheltenham, 381 pp.

Kelly, P.D., Chu, F., Woods, I.G., Ngo-Hazelett, P., Cardozo, T., Huang, H., Kimm, F., Liao, L., Yan, Y.L., Zhou, Y., Johnson, S.L., Abagyan, R., Schier, A.F., Postlethwait, J.H., Talbot, W.S., 2000. Genetic linkage mapping of zebrafish genes and ESTs. Genome Research 10, 558-567.

Khoo, S.K., Ozaki, A., Nakamura, F., Arakawa, T., Ishimoto, S., Nickolov, R., Sakamoto, T., Akutsu, T., Mochizuki, M., Denda, I., Okamoto, N., 2004. Identification of a novel chromosomal region associated with infectious hematopoietic necrosis (IHN) resistance in rainbow trout *Oncorhynchus mykiss*. Fish Pathology 39, 95-101.

Kijas, J.M., Fowler, J.C., Garbett, C.A., Thomas, M.R., 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. Biotechniques 16, 656-660, 662.

Kimura, M., Crow, J.F., 1964. The number of alleles that can be maintained in a finite population. Genetics 49, 725-738.

King, J., Armstead, I.P., Donnison, I.S., Thomas, H.M., Jones, R.N., Kearsey, M.J., Roberts, L.A., Thomas, A., Morgan, W.G., King, I.P., 2002. Physical and genetic mapping in the grasses *Lolium perenne* and *Festuca pratensis*. Genetics 161, 315-324.

King, P.A., McGrath, D., Britton, W., 1990. The use of artificial substrates in monitoring mussel (*Mytilus edulis* L.) settlement on an exposed rocky shore in the west coast of Ireland. Journal of the Marine Biological Association of the United Kingdom 70, 371-380.

Kinnison, M.T., Bentzen, P., Unwin, M.J., Quinn, T.P., 2002. Reconstructing recent divergence: evaluating nonequilibrium population structure in New Zealand chinook salmon. Molecular Ecology 11, 739-754.

Knapik, E.W., Goodman, A., Ekker, M., Chevrette, M., Delgado, J., Neuhauss, S., Shimoda, N., Driever, W., Fishman, M.C., Jacob, H.J., 1998. A microsatellite genetic linkage map for zebrafish (*Danio rerio*). Nature Genetics 18, 338-343.

Kocher, T.D., Lee, W.J., Sobolewska, H., Penman, D., McAndrew, B., 1998. A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). Genetics 148, 1225-1232.

Kornberg, A., 1980. In: Freeman, W.H. (Ed.), DNA replication, San Francisco.

Kosambi, D.D., 1944. The estimation of map distances from recombination values. Annals of Eugenics 12, 172-175.

Krauss, S.L., 1999. Complete exclusion of nonsires in an analysis of paternity in a natural plant population using amplified fragment length polymorphism (AFLP). Molecular Ecology 8, 217-226.

Lai, Y., Sun, F., 2003. The relationship between microsatellite slippage mutation rate and the number of repeat units. Molecular Biology and Evolution 20, 2123-2131.

Lallias, D., Taris, N., Boudry, P., Bonhomme, F., Lapègue, S., in prep. Variance in reproductive success of the flat oyster *Ostrea edulis* L. assessed by microsatellite-based parentage analyses in natural populations and under experimental conditions.

Lama, A., Montes, J., 1993. Influence of depth of culture in the infection of the European flat oyster (*Ostrea edulis* L.) by *Bonamia ostreae*. Bulletin of the European Association of Fish Pathologists 13, 17-20.

Lander, E.S., Botstein, D., 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121, 185-199.

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E., Newburg, L., 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1, 174-181.

Lane, D.J.W., Beaumont, A.R., Hunter, J.R., 1985. Byssus drifting and the drifting threads of young postlarval mussel *Mytilus edulis*. Marine Biology 84, 301-308.

Langdon, C., Jacobson, D., Matson, S., Evans, F., 2003. Improvement of yields of Pacific oysters through genetic selection. Journal of Shellfish Research 22, 327.

Lannan, J.E., 1972. Estimation of heritability and predicting response to selection for the Pacific oyster, *Crassostrea gigas*. Proceedings of the National Shellfisheries Association 62, 62-66.

Lannan, J.E., 1980a. Broodstock management of *Crassostrea gigas*: III. Selective breeding for improved larval survival. Aquaculture 21, 347-351.

Lannan, J.E., 1980b. Broodstock management of *Crassostrea gigas*: I. Genetic and environmental variation in survival in the larval rearing system. Aquaculture 21, 323-336.

Launey, S., 1998. Marqueurs microsatellites chez l'huître plate *Ostrea edulis* L.: caractérisation et applications à un programme de sélection pour une résistance au parasite *Bonamia ostreae* et à l'etude de populations naturelles. Thèse de doctorat. Institut national agronomique Paris Grignon, pp. 305 p.

Launey, S., Hedgecock, D., 2001. High genetic load in the Pacific oyster *Crassostrea gigas*. Genetics 159, 255-265.

Launey, S., Barre, M., Gerard, A., Naciri-Graven, Y., 2001. Population bottleneck and effective size in *Bonamia ostreae*-resistant populations of *Ostrea edulis* as inferred by microsatellite markers. Genetical Research 78, 259-270.

Launey, S., Ledu, C., Boudry, P., Bonhomme, F., Naciri-Graven, Y., 2002. Geographic structure in the European flat oyster (*Ostrea edulis* L.) as revealed by microsatellite polymorphism. Journal of Heredity 93, 331-351.

Le Dantec, J., Marteil, L., 1976. La reproduction des huitres. Revue des Travaux de l'Institut des Peches marines 40, 233-256.

Lee, B.Y., Penman, D.J., Kocher, T.D., 2003. Identification of a sex-determining region in Nile tilapia (*Oreochromis niloticus*) using bulked segregant analysis. Animal Genetics 34, 379-383.

Lehnert, S.A., Wilson, K.J., Byrne, K., Moore, S.S., 1999. Tissue-Specific Expressed Sequence Tags from the black tiger shrimp *Penaeus monodon*. Marine Biotechnology 1, 465-476.

Levinson, G., Gutman, G.A., 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Molecular Biology and Evolution 4, 203-221.

Lewis, J.R., 1977. The role of physical and biological factors in the distribution and stability of rocky shore communities. In: B.F. Keegan, P.O.C.P.J.B. (Ed.), Proceedings of the 11th European Marine Biology Symposium. Oxford: Pergamon Press, Galway, Ireland, 5-11 October 1976, pp. 417-424.

Li, G., Hedgecock, D., 1998. Genetic heterogeneity, detected by PCR-SSCP, among larval samples of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of large variance in reproductive success. Canadian Journal of Fisheries and Aquatic Sciences 55, 1025-1033.

Li, L., Guo, X., 2004. AFLP-based genetic linkage maps of the pacific oyster *Crassostrea gigas* Thunberg. Marine Biotechnology 6, 26-36.

Li, L., Xiang, J., Liu, X., Zhang, Y., Dong, B., Zhang, X., 2005. Construction of AFLP-based genetic linkage map for Zhikong scallop, *Chlamys farreri* Jones et Preston and mapping of sex-linked markers. Aquaculture 245, 63-73.

Li, Q., Yu, H., Yu, R., 2006c. Genetic variability assessed by microsatellites in cultured populations of the Pacific oyster (*Crassostrea gigas*) in China. Aquaculture 259, 95-102.

Li, Y., Dierens, L., Byrne, K., Miggiano, E., Lehnert, S., Preston, N., Lyons, R., 2006a. QTL detection of production traits for the Kuruma prawn *Penaeus japonicus* (Bate) using AFLP markers. Aquaculture 258, 198-210.

Li, Y., Byrne, K., Miggiano, E., Whan, V., Moore, S., Keys, S., Crocos, P., Preston, N., Lehnert, S., 2003. Genetic mapping of the kuruma prawn *Penaeus japonicus* using AFLP markers. Aquaculture 219, 143-156.

Li, Y.C., Korol, A.B., Fahima, T., Beiles, A., Nevo, E., 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Molecular Ecology 11, 2453-2465.

Li, Z., Li, J., Wang, Q., He, Y., Liu, P., 2006b. AFLP-based genetic linkage map of marine shrimp *Penaeus* (*Fenneropenaeus*) *chinensis*. Aquaculture 261, 463-472.

Lincoln, S.E., Lander, E.S., 1992. Systematic detection of errors in genetic linkage data. Genomics 14, 604-610.

Lincoln, S.E., Daly, A., Lander, E.S., 1992. Mapping genes controlling quantitative traits with MAPMAKER/QTL. Whitehead Institute Technical Report, 2nd ed.

Litt, M., Luty, J.A., 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetics 44, 397-401.

Liu, X., Liu, X., Guo, X., Gao, Q., Zhao, H., Zhang, G., 2006. A preliminary genetic linkage map of the Pacific abalone *Haliotis discus hannai* Ino. Marine Biotechnology 8, 1-12.

Liu, Z., Karsi, A., Li, P., Cao, D., Dunham, R., 2003. An AFLP-based genetic linkage map of channel catfish (*Ictalurus punctatus*) constructed by using an interspecific hybrid resource family. Genetics 165, 687-694.

Liu, Z.J., Cordes, J.F., 2004. DNA marker technologies and their applications in aquaculture genetics. Aquaculture 238, 1-37.

Liu, Z.J., Li, P., Argue, B.J., Dunham, R.A., 1999. Random amplified polymorphic DNA markers: usefulness for gene mapping and analysis of genetic variation of catfish. Aquaculture 174, 59-68.

Longwell, A.C., Stiles, S.S., 1973. Oyster genetics and the probable future role of genetics in aquaculture. Malacological Review 6, 151-177.

Losee, E., 1978. Influence of heredity on larval and spat growth in *Crassostrea virginica*. Proceedings of the 9th Annual meeting of the World Mariculture Society 9, 101-107.

Lunt, D.H., Hutchinson, W.F., Carvalho, G.R., 1999. An efficient method for PCR-based identification of microsatellite arrays (PIMA). Molecular Ecology 8, 891-892.

Lutz, R.A., Kennish, M.J., 1992. Ecology and morphology of larval and early larval postlarval mussels. In: Gosling, E.M. (Ed.), The mussel *Mytilus*: ecology, physiology, genetics and culture, (ed.). Elsevier Science Publ., Amsterdam, pp. 53-85.

Lynch, M., Walsh, B., 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, 980 pp.

Lynch, S.A., Wylde, S., Armitage, D.V., Mulcahy, M.F., Culloty, S.C., 2005. The susceptibility of young, prespawning oysters, *Ostrea edulis*, to *Bonamia ostreae*. Journal of Shellfish Research 24, 664.

Maheswaran, M., Subudhi, P.K., Nandi, S., Xu, J.C., Parco, A., Yang, D.C., Huang, N., 1997. Polymorphism, distribution, and segregation of AFLP markers in a doubled haploid rice population. Theoretical and Applied Genetics 94, 39-45.

Malecha, S.R., Masuno, S., Onizuka, D., 1984. The feasibility of measuring the heritability of growth pattern variation in juvenile freshwater prawns, *Macrobrachium rosenbergii* (de Man). Aquaculture 38, 347-363.

Mallet, A.L., 1988. Genetic analysis of growth and viability in the blue mussel, *Mytilus edulis*. Journal of Shellfish Research 7, 168.

Mallet, A.L., Freeman, K.R., Dickie, L.M., 1986. The genetics of production characters in the blue mussel *Mytilus edulis*. I. A preliminary analysis. Aquaculture 57, 133-140.

Manly, K.F., Olson, J.M., 1999. Overview of QTL mapping software and introduction to map manager QT. Mammalian Genome 10, 327-334.

Marshall, T.C., Slate, J., Kruuk, L.E., Pemberton, J.M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7, 639-655.

Martin, A.-G., Gérard, A., Cochennec, N., Langlade, A., 1993. Selecting flat oysters, *Ostrea edulis*, for survival against the parasite *Bonamia ostreae*: assessment of the resistance of a first selected generation. Production, Environment and Quality. Bordeaux Aquaculture 92, European Aquaculture Society, Ghent, Belgium, pp. 547-554. Special publication No. 518.

Martinez-Lage, A., Gonzalez-Tizon, A., Mendez, J., 1995. Chromosomal markers in three species of the genus *Mytilus* (Mollusca: Bivalvia). Heredity 74, 369-375.

Martinez-Lage, A., Gonzalez-Tizon, A., Mendez, J., 1996. Chromosomes differences between European mussel populations (genus *Mytilus*). Caryologia 49, 343-355.

Martinez, O., Curnow, R.N., 1992. Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theoretical and Applied Genetics 85, 480-488.

Martyniuk, C.J., Perry, G.M.L., Mogahadam, H.K., Ferguson, M.M., Danzmann, R.G., 2003. The genetic architecture of correlations among growth-related traits and male age at maturation in rainbow trout. Journal of Fish Biology 63, 746-764.

Matthes, M.C., Daly, A., Edwards, K.J., 1998. Amplified fragment length polymorphism (AFLP). In: Angela Karp, P.G.I.a.D.S.I. (Ed.), Molecular tools for sreening biodiversity. Chapman & Hall, London, pp. 183-192.

Maughan, P.J., Maroof, M.A.S., Buss, G.R., Huestis, G.M., 1996. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. Theoretical and Applied Genetics 93, 392-401.

May, B., Johnson, K.R., 1989. Composite linkage map of Salmonid fishes (*Salvelinus*, *Salmo*, *Oncorhynchus*). In: O' Brien, S. J. (Ed.), Genetic maps: locus maps of complex genomes, volume 4, pp. 151-159.

McGoldrick, D.J., Hedgecock, D., 1997. Fixation, segregation and linkage of allozyme loci in inbred families of the Pacific oyster *Crassostrea gigas* (Thunberg): implications for the causes of inbreeding depression. Genetics 146, 321-334.

McGoldrick, D.J., Hedgecock, D., English, L.J., Baoprasertkul, P., Ward, R.D., 2000. The transmission of microsatellite alleles in Australian and North American stocks of the Pacific oyster (*Crassostrea gigas*): Selection and null alleles. Journal of Shellfish Research 19, 779-788.

McGrath, D., King, P.A., Gosling, E.M., 1988. Evidence for the direct settlement of *Mytilus edulis* L. larvae on adult mussel beds. Marine Ecology Progress Series 47, 103-106.

McKay, L.R., Ihssen, P.E., Friars, G.W., 1986. Genetic parameters of growth in rainbow trout, *Salmo gairdneri*, as a function of age and maturity. Aquaculture 58, 241-254.

Mialhe, E., Bachere, E., Chagot, D., Grizel, H., 1988. Isolation and purification of the protozoan *Bonamia ostreae* (Pichot et al. 1980), a parasite affecting the flat oyster *Ostrea edulis* L. Aquaculture 71, 293-299.

Michelmore, R.W., Paran, I., Kesseli, R.V., 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proceedings of the National Academy of Sciences of the USA 88, 9828-9832.

Miller, P.S., Hedrick, P.W., 2001. Purging of inbreeding depression and fitness decline in bottlenecked populations of *Drosophila melanogaster*. Journal of Evolutionary Biology 14, 595-601.

Moen, T., Fjalestad, K.T., Munck, H., Gomez-Raya, L., 2004a. A multistage testing strategy for detection of quantitative trait loci affecting disease resistance in Atlantic salmon. Genetics 167, 851-858.

Moen, T., Hoyheim, B., Munck, H., Gomez-Raya, L., 2004b. A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. Animal Genetics 35, 81-92.

Moen, T., Agresti, J.J., Cnaani, A., Moses, H., Famula, T.R., Hulata, G., Gall, G.A.E., May, B., 2004c. A genome scan of a four-way tilapia cross supports the existence of a quantitative trait locus for cold tolerance on linkage group 23. Aquaculture Research 35, 893-904.

Montes, J., Ferro-Soto, B., Conchas, R.F., Guerra, A., 2003. Determining culture strategies in populations of the European flat oyster, *Ostrea edulis*, affected by bonamiosis. Aquaculture 220, 175-182.

Moore, S.S., Sargeant, L.L., King, T.J., Mattick, J.S., Georges, M., Hetzel, D.J., 1991. The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. Genomics 10, 654-660.

Moore, S.S., Whan, V., Davis, G.P., Byrne, K., Hetzel, D.J.S., Preston, N., 1999. The development and application of genetic markers for the Kuruma prawn *Penaeus japonicus*. Aquaculture 173, 19-32.

Moreno, F.J., Hernandez-Verdun, D., Masson, C., Bouteille, M., 1985. Silver staining of the Nucleolar Organizer Regions (NORs) on lowicryl and cryo-unltrathin sections. The Journal of Histochemestry and Cytochemestry 33, 389-399.

Morgan, T.S., Rogers, A.D., 2001. Specificity and sensitivity of microsatellite markers for the identification of larvae. Marine Biology 139, 967-973.

Morgan, T.S., Rogers, A.D., Iyengar, A., 2000. Novel microsatellite markers for the European oyster *Ostrea edulis*. Molecular Ecology 9, 495-497.

Morin, P.A., Wallis, J., Moore, J.J., Woodruff, D.S., 1994. Paternity exclusion in a community of wild chimpanzees using hypervariable simple sequence repeats. Molecular Ecology 3, 469-477.

Morizot, D.C., Slaugenhaupt, S.A., Kallman, K.D., Chakravarti, A., 1991. Genetic linkage map of fishes of the genus *Xiphophorus* (Teleostei: Poeciliidae). Genetics 127, 399-410.

Naciri-Graven, Y., Haure, J., Gerard, A., Baud, J.-P., 1999. Comparative growth of *Bonamia ostreae* resistant and wild flat oyster *Ostrea edulis* in an intensive system: II. Second year of the experiment. Aquaculture 171, 195-208.

Naciri-Graven, Y., Martin, A.-G., Baud, J.-P., Renault, T., Gerard, A., 1998. Selecting the flat oyster *Ostrea edulis* (L.) for survival when infected with the parasite *Bonamia ostreae*. Journal of Experimental Marine Biology and Ecology 224, 91-107.

Naciri-Graven, Y., Launey, S., Lebayon, N., Gerard, A., Baud, J.P., 2000. Influence of parentage upon growth in *Ostrea edulis*: evidence for inbreeding depression. Genetical Research 76, 159-168.

Naciri, Y., Vigouroux, Y., Dallas, J., Desmarais, E., Delsert, C., Bonhomme, F., 1995. Identification and inheritance of (GA/TC)n and (AC/GT)n repeats in the European flat oyster *Ostrea edulis* (L.). Molecular Marine Biology and Biotechnology 4, 83-89.

Nakamura, K., Ozaki, A., Akutsu, T., Iwai, K., Sakamoto, T., Yoshizaki, G., Okamoto, N., 2001. Genetic mapping of the dominant albino locus in rainbow trout (*Oncorhynchus mykiss*). Molecular Genetics and Genomics 265, 687-693.

Naruse, K., Fukamachi, S., Mitani, H., Kondo, M., Matsuoka, T., Kondo, S., Hanamura, N., Morita, Y., Hasegawa, K., Nishigaki, R., Shimada, A., Wada, H., Kusakabe, T., Suzuki, N., Kinoshita, M., Kanamori, A., Terado, T., Kimura, H., Nonaka, M., Shima, A., 2000. A detailed linkage map of medaka, *Oryzias latipes*: comparative genomics and genome evolution. Genetics 154, 1773-1784.

Nei, M., 1978. Estimation of Average Heterozygosity and genetic distance from a small number of individuals. Genetics 89, 583-590.

Neild, R., 1995. The English, the French and the oyster. Quiller Press, London.

Newkirk, G.F., 1980. Review of the genetics and the potential for selective breeding of commercially important bivalves. Aquaculture 19, 209-228.

Newkirk, G.F., Haley, L.E., 1982. Progress in selection for growth rate in the European flat oyster *Ostrea edulis*. Marine Ecology Progress Series 10, 77-79.

Newkirk, G.F., Haley, L.E., 1983. Selection for growth rate in the European oyster, *Ostrea edulis*: Response of second generation groups. Aquaculture 33, 149-155.

Newkirk, G.F., Haley, L.E., Waugh, D.L., Doyle, R., 1977. Genetics of larvae and spat growth rate in the oyster *Crassostrea virginica*. Marine Biology 41, 49-52.

Nichols, K.M., Wheeler, P.A., Thorgaard, G.H., 2004. Quantitative trait loci analyses for meristic traits in *Oncorhynchus mykiss*. Environmental biology of fishes 69, 317-331.

Nichols, K.M., Young, W.P., Danzmann, R.G., Robison, B.D., Rexroad, C., Noakes, M., Phillips, R.B., Bentzen, P., Spies, I., Knudsen, K., Allendorf, F.W., Cunningham, B.M., Brunelli, J., Zhang, H., Ristow, S., Drew, R., Brown, K.H., Wheeler, P.A., Thorgaard, G.H., 2003. A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*). Animal Genetics 34, 102-115.

Nilsson, J., 1990. Heritability estimates of growth-related traits in Arctic charr (*Salvelinus alpinus*). Aquaculture 84, 211-217.

O'Malley, K.G., Sakamoto, T., Danzmann, R.G., Ferguson, M.M., 2003. Quantitative trait loci for spawning date and body weight in rainbow trout: testing for conserved effects across ancestrally duplicated chromosomes. Journal of Heredity 94, 273-284.

O'Neill, G., Culloty, S.C., Mulcahy, M.F., 1998. The effectiveness of two routine doagnostic techniques for the detection of the protozoan parasite, *Bonamia ostreae* (Pichot et al. 1980). Bulletin of the European Association of Fish Pathologists 18, 117-120.

O'Reilly, P.T., Herbinger, C., Wright, J.M., 1998. Analysis of parentage determination in Atlantic salmon (*Salmo salar*) using microsatellites. Animal Genetics 29, 363-370.

Ohara, E., Nishimura, T., Nagakura, Y., Sakamoto, T., Mushiake, K., Okamoto, N., 2005. Genetic linkage maps of two yellowtails (*Seriola quinqueradiata* and *Seriola lalandi*). Aquaculture 244, 41-48.

Ostrander, E.A., Jong, P.M., Rine, J., Duyk, G., 1992. Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. Proceedings of the National Academy of Sciences of the USA 89, 3419-3423.

Ota, T., Kimura, M., 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genetical Research 22, 201-204.

Otsen, M., den Bieman, M., Kuiper, M.T., Pravenec, M., Kren, V., Kurtz, T.W., Jacob, H.J., Lankhorst, A., van Zutphen, B.F., 1996. Use of AFLP markers for gene mapping and QTL detection in the rat. Genomics 37, 289-294.

Ozaki, A., Sakamoto, T., Khoo, S., Nakamura, K., Coimbra, M.R., Akutsu, T., Okamoto, N., 2001. Quantitative trait loci (QTLs) associated with resistance/susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*). Molecular Genetics and Genomics 265, 23-31.

Paetkau, D., 1999. Microsatellites obtained using strand extension: an enrichment protocol. Biotechniques 26, 690-692, 694-697.

Palti, Y., Parsons, J.E., Thorgaard, G.H., 1999. Identification of candidate DNA markers associated with IHN virus resistance in backcrosses of rainbow (*Oncorhynchus mykiss*) and cutthroat trout (*O. clarki*). Aquaculture 173, 81-94.

Pemberton, J.M., Slate, J., Bancroft, D.R., Barrett, J.A., 1995. Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. Molecular Ecology 4, 249-252.

Perez-Rostro, C.I., Ibarra, A.M., 2003. Quantitative genetic parameter estimates for size and growth rate traits in Pacific white shrimp, *Penaeus vannamei* (Boone 1931) when reared indoors. Aquaculture Research 34, 543-553.

Perez, F., Erazo, C., Zhinaula, M., Volckaert, F., Calderon, J., 2004. A sex-specific linkage map of the white shrimp *Penaeus* (*Litopenaeus*) *vannamei* based on AFLP markers. Aquaculture 242, 105-118.

Perry, G.M., Ferguson, M.M., Danzmann, R.G., 2003. Effects of genetic sex and genomic background on epistasis in rainbow trout (*Oncorhynchus mykiss*). Genetica 119, 35-50.

Perry, G.M., Danzmann, R.G., Ferguson, M.M., Gibson, J.P., 2001. Quantitative trait loci for upper thermal tolerance in outbred strains of rainbow trout (*Oncorhynchus mykiss*). Heredity 86, 333-341.

Phillips, H.A., Thompson, A.D., Kozman, H.M., Sutherland, G.R., Mulley, J.C., 1993. A microsatellite marker within the duplicated D16S79 locus has a null allele; significance for linkage mapping. Cytogenetics and Cell Genetics 64, 131-132.

Pichot, Y., Comps, M., Tigé, G., Grizel, H., Rabouin, M.A., 1980. Recherches sur *Bonamia ostreae* gen. n., sp. n., parasite nouveau de l'huître plate *Ostrea edulis* L. Revue des Travaux de l'Institut des Pêches Maritimes 43, 131-140.

Pillay, T.V.R., 1990. Genetic selection and hybridization. In: Pillay, T.V.R. (Ed.), Aquaculture: Principles and Practices. Fishing News Books, Oxford, pp. 164-173.

Plomion, C., O'Malley, D.M., 1996. Recombination rate differences for pollen parents and seed parents in *Pinus pinaster*. Heredity 77, 341-350.

Poompuang, S., Na-Nakorn, U., 2004. A preliminary genetic map of walking catfish (*Clarias macrocephalus*). Aquaculture 232, 195-203.

Postlethwait, J.H., Johnson, S.L., Midson, C.N., Talbot, W.S., Gates, M., Ballinger, E.W., Africa, D., Andrews, R., Carl, T., Eisen, J.S., et al., 1994. A genetic linkage map for the zebrafish. Science 264, 699-703.

Presa, P., Pérez, M., Diz, A.P., 2002. Polymorphic microsatellite markers for blue mussels (*Mytilus* spp.). Conservation genetics 3, 441-443.

Primmer, C.R., Moller, A.P., Ellegren, H., 1996. A wide-range survey of cross-species microsatellite amplification in birds. Molecular Ecology 5, 365-378.

Queller, D.C., Strassmann, J.E., Hughes, C.R., 1993. Microsatellites and kinship. Trends in Ecology & Evolution 8, 285-288.

Quero, J.C., Vayne, J.J., 1998. Les fruits de la mer et plantes marines des pêches françaises. Les encyclopédies du naturaliste, Delachaux Niestlé.

Questiau, S., Eybert, M.-C., Taberlet, P., 1999. Amplified fragment length polymorphism (AFLP) markers reveal extra-pair parentage in a bird species: the bluethroat (*Luscinia svecica*). Molecular Ecology 8, 1331-1339.

Rassmann, K., Schlotterer, C., Tautz, D., 1991. Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. Electrophoresis 12, 113-118.

Rawson, P.D., Hilbish, T.J., 1990. Heritability of juvenile growth for the hard clam *Mercenaria mercenaria*. Marine Biology 105, 429-436.

Refstie, T., 1980. Genetic and environmental sources of variation in body weight and length of rainbow trout fingerlings. Aquaculture 19, 351-357.

Refstie, T., Steine, T.A., 1978. Selection experiments with salmon: III. Genetic and environmental sources of variation in length and weight of Atlantic salmon in the freshwater phase. Aquaculture 14, 221-234.

Richardson, C.A., Seed, R., 1990. Predictions of mussel (*Mytilus edulis*) biomass on an offshore platform from single population samples. Biofouling 2, 289-297.

Ritter, E., Gebhardt, C., Salamini, F., 1990. Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. Genetics 125, 645-654.

Robert, R., Borel, M., Pichot, Y., Trut, G., 1991. Growth and mortality of the European oyster *Ostrea edulis* in the Bay of Arcachon (France). Aquatic Living Resources 4, 265-274.

Robison, B.D., Wheeler, P.A., Sundin, K., Sikka, P., Thorgaard, G.H., 2001. Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). Journal of Heredity 92, 16-22.

Rodríguez-Juíz, A.M., Torrado, M., Méndez, J., 1996. Genome-size variation in bivalve molluscs determined by flow cytometry. Marine Biology 126, 489-497.

Rodriguez, M.F., LaPatra, S., Williams, S., Famula, T., May, B., 2004. Genetic markers associated with resistance to infectious hematopoietic necrosis in rainbow and steelhead trout (*Oncorhynchus mykiss*) backcrosses. Aquaculture 241, 93-115.

Rogan, E., Culloty, S.C., Cross, T.F., Mulcahy, M.F., 1991. The detection of *Bonamia ostreae* (Pichot et al. 1980) in frozen oysters (*Ostrea edulis* L.) and the effect on the parasite condition. Aquaculture 97, 311-315.

Rohrer, G.A., Alexander, L.J., Hu, Z., Smith, T.P., Keele, J.W., Beattie, C.W., 1996. A comprehensive map of the porcine genome. Genome Research 6, 371-391.

Rose, C., Paynter, K., Reece, K., Hare, M., 2005. Isolation by distance in the eastern oyster, *Crassostrea virginica*, in Chesapeake Bay. Journal of Shellfish Research 24, 673.

Ruzzante, D.E., 1986. Selection for growth rate in the European oyster *Ostrea edulis*: a multivariate approach., M.Sc. thesis. Dalhousie University, Halifax.

Saavedra, C., Zapata, C., Alvarez, G., 1995. Geographical patterns of variability at allozyme loci in the European oyster *Ostrea edulis*. Marine Biology 122, 95-104.

Saavedra, C., Zapata, C., Guerra, A., Alvarez, G., 1993. Allozyme variation in European populations of the oyster *Ostrea edulis*. Marine Biology 115, 85-95.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A., 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239, 487-491.

Sakamoto, T., Danzmann, R.G., Okamoto, N., Ferguson, M.M., Ihssen, P.E., 1999. Linkage analysis of quantitative trait loci associated with spawning time in rainbow trout (*Oncorhynchus mykiss*). Aquaculture 173, 33-43.

Sakamoto, T., Danzmann, R.G., Gharbi, K., Howard, P., Ozaki, A., Khoo, S.K., Woram, R.A., Okamoto, N., Ferguson, M.M., Holm, L.E., Guyomard, R., Hoyheim, B., 2000. A

microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. Genetics 155, 1331-1345.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press.

Satagopan, J.M., Yandell, B.S., Newton, M.A., Osborn, T.C., 1996. A bayesian approach to detect quantitative trait loci using Markov chain Monte Carlo. Genetics 144, 805-816.

Sax, K., 1923. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8, 552-560.

Schlotterer, C., Tautz, D., 1992. Slippage synthesis of simple sequence DNA. Nucleic Acids Research 20, 211-215.

Seaton, G., Haley, C.S., Knott, S.A., Kearsey, M., Visscher, P.M., 2002. QTL Express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics 18, 339-340.

Seed, R., 1968. Factors influencing shell shape in the mussel *Mytilus edulis*. Journal of the Marine Biological Association of the United Kingdom 48, 561-584.

Seed, R., Suchanek, T.H., 1992. Population and community ecology of *Mytilus*. In: Gosling, E.M. (Ed.), The mussel *Mytilus*: ecology, physiology, genetics and culture. Elsevier Science Publ., Amsterdam, pp. 87-169.

Shah, M.M., Hassan, A., 2005. Distribution of genes and recombination on wheat homoeologous group 6 chromosomes: a synthesis of available information. Molecular Breeding 15, 45-53.

Sharbel, T.F., 1999. Amplified Fragment Length Polymorphisms: a non-random PCR-based technique for multilocus sampling. In: Epplen, J.T., Lubjuhn, T. (Ed.), DNA profiling and DNA fingerprinting. Birkhauser Verlag, Basel, pp. 177-194.

Sheridan, A.K., 1997. Genetic improvement of oyster production--a critique. Aquaculture 153, 165-179.

Shimoda, N., Knapik, E.W., Ziniti, J., Sim, C., Yamada, E., Kaplan, S., Jackson, D., de Sauvage, F., Jacob, H., Fishman, M.C., 1999. Zebrafish genetic map with 2000 microsatellite markers. Genomics 58, 219-232.

Shriver, M.D., Jin, L., Chakraborty, R., Boerwinkle, E., 1993. VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. Genetics 134, 983-993.

Sillanpaa, M.J., Arjas, E., 1998. Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. Genetics 148, 1373-1388.

Slate, J., Marshall, T., Pemberton, J., 2000. A retrospective assessment of the accuracy of the paternity inference program CERVUS. Molecular Ecology 9, 801-808.

Slettan, A., Olsaker, I., Lie, O., 1997. Segregation studies and linkage analysis of Atlantic salmon microsatellites using haploid genetics. Heredity 78 (Pt 6), 620-627.

Sobolewska, H., Beaumont, A.R., 2005. Genetic variation at microsatellite loci in northern populations of the European flat oyster (*Ostrea edulis*). Journal of the Marine Biological Association of the United Kingdom 85, 955-960.

Sobolewska, H., Beaumont, A.R., Hamilton, A., 2001. Dinucleotide microsatellites isolated from the European flat oyster, *Ostrea edulis*. Molecular Ecology Notes 1, 79-80.

Somorjai, I.M., Danzmann, R.G., Ferguson, M.M., 2003. Distribution of temperature tolerance quantitative trait loci in Arctic charr (*Salvelinus alpinus*) and inferred homologies in rainbow trout (*Oncorhynchus mykiss*). Genetics 165, 1443-1456.

Stam, P., 1993. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. The plant journal 3, 739-744.

Stromgren, T., Nielsen, M.V., 1989. Heritability of growth in larvae and juveniles of *Mytilus edulis*. Aquaculture 80, 1-6.

Suchanek, T.H., 1978. The ecology of *Mytilus edulis* L. in exposed rocky intertidal communities. Journal of Experimental Marine Biology and Ecology 31, 105-120.

Suchanek, T.H., 1985. Mussels and their role in structuring rocky shore communities. In: Seed, P.G.M.R. (Ed.), The Ecology of Rocky Coasts: essays presented to J.R. Lewis, D.Sc., pp. 70-96.

Suiter, K.A., Wendel, J.F., Case, J.S., 1983. LINKAGE-1: a PASCAL computer program for the detection and analysis of genetic linkage. The journal of heredity 74, 203-204.

Sun, X., Liang, L., 2004. A genetic linkage map of common carp (*Cyprinus carpio* L.) and mapping of a locus associated with cold tolerance. Aquaculture 238, 165-172.

Sunnucks, P., 2000. Efficient genetic markers for population biology. Trends in Ecology and Evolution 15, 199-203.

Tanguy, A., Guo, X., Ford, S.E., 2004. Discovery of genes expressed in response to *Perkinsus marinus* challenge in Eastern (*Crassostrea virginica*) and Pacific (*C. gigas*) oysters. Gene 338, 121-131.

Tanguy, A., Boutet, I., Laroche, J., Moraga, D., 2005. Molecular identification and expression study of differentially regulated genes in the Pacific oyster *Crassostrea gigas* in response to pesticide exposure. The FEBS Journal 272, 390-403.

Tanksley, S.D., 1993. Mapping polygenes. Annual Review of Genetics 27, 205-233.

Tanksley, S.D., Ganal, M.W., Prince, J.P., de-Vicente, M.C., Bonierbale, M.W., Broun, P., Fulton, T.M., Giovannoni, J.J., Grandillo, S., Martin, G.B., Messeguer, R., Miller, J.C., Miller, L., Paterson, A.H., Pineda, O., Roder, M.S., Wing, R.A., Wu, W., Young, N.D., 1992. High density molecular linkage maps of the tomato and potato genomes. Genetics 132, 1141-1160.

Taris, N., Boudry, P., Bonhomme, F., Lapègue, S. Limited evidence of sweepstakes reproductive success in temporal cohorts of European flat oysters juveniles. Submitted to Marine Biology.

Thiriot-Quiévreux, C., 1984a. Chromosome analysis of three species of *Mytilus* (Bivalvia: Mytilidae). Marine Biology Letter 5, 265-273.

Thiriot-Quiévreux, C., 1984b. Analyse comparée des caryotypes d'ostreidae (Bivalvia). Cahiers de Biologie Marine 25, 407-418.

Thiriot-Quiévreux, C., 2002. Review of the literature on bivalve cytogenetics in the last ten years. Cahiers de Biologie Marine 43, 17-26.

Thiriot-Quiéveux, C., Ayraud, N., 1982. Les caryotypes de quelques espèces de bivalves et de gastéropodes marins. Marine Biology 70, 165-172.

Thiriot-Quiévreux, C., Insua, A., 1992. Nucleolar organiser region variation in the chromosomes of three oyster species. Journal of Experimental Marine Biology and Ecology 157, 33-40.

Thoday, J.M., 1961. Location of polygenes. Nature 191, 368-370.

Thompson, R.J., 1979. Fecundity and reproductive effort of the blue mussel (*Mytilus edulis*), the sea urchin (*Strongylocentrotus droebachiensis*) and the snow crab (*Chionectes opilio*) from populations in Nova Scotia and Newfoundland. Journal of the Fisheries Research Board of Canada 36, 955-964.

Thorgaard, G.H., Allendorf, F.W., Knudsen, K.L., 1983. Gene-centromere mapping in rainbow trout: high interference over long map distances. Genetics 103, 771-783.

Tigé, G., Grizel, H., 1984. Essai de contamination d'*Ostrea edulis* Linné par *Bonamia ostreae* (Pichot et al., 1979) en rivière de Crach (Morbihan). Revue des Travaux de l'Institut des Pêches Maritimes 328, 3-13.

Tinker, N.A., Mather, D.E., 1995. MQTL: software for simplified composite interval mapping of QTL in multiple environments. Journal of Quantitative Trait Loci 1.

Toro, J.E., Newkirk, G.F., 1990. Divergent selection for growth rate in the European oyster *Ostrea edulis*: response to selection and estimation of genetic parameters. Marine Ecology Progress Series 62, 219-227.

Toro, J.E., Paredes, L.I., 1996. Heritability estimates of larval shell length in the Chilean blue mussel *Mytilus chilensis*, under different food densities. Aquatic Living Resources 9, 347-350.

Toro, J.E., Aguila, P., Vergara, A.M., Newkirk, G.F., 1994. Realized heritability estimates for growth from data on tagged Chilean native oyster (*Ostrea chilensis*). World Aquaculture 25, 29-30.

Toro, J.E., Sanhueza, M.A., Winter, J.E., Aguila, P., Vergara, A.M., 1995. Selection response and heritability estimates for growth in the Chilean oyster *Ostrea chilensis* (Philippi, 1845). Journal of Shellfish Research 14, 87-92.

Toth, G., Gaspari, Z., Jurka, J., 2000. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Research 10, 967-981.

Travis, S.E., Maschinski, J., Keim, P., 1996. An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically endangered plant, using AFLP markers. Molecular Ecology 5, 735-745.

Treco, D., Arnheim, N., 1986. The evolutionarily conserved repetitive sequence d(TG.AC)n promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. Molecular and Cell Biology 6, 3934-3947.

Tsuchiya, M., 1980. Biodeposit production by the mussel *Mytilus edulis* L. on rocky shores. Journal of Experimental Marine Biology and Ecology 47, 203-222.

Tsuchiya, M., Nishihira, M., 1986. Islands of *Mytilus edulis* as a habitat for small intertidal animals: effect of *Mytilus* age structure on the species composition of the associated fauna and community organization. Marine Ecology Progress Series 31, 171-178.

Uimari, P., Thaller, G., Hoeschele, I., 1996. The use of multiple markers in a Bayesian method for mapping quantitative trait loci. Genetics 143, 1831-1842.

Utting, S.D., Spencer, B.E., 1991. The hatchery culture of bivalve mollusc larvae and juveniles, Laboratory leaflet number 68. Ministry of agriculture, fisheries and food directorate of fisheries research, Lowestoft, pp. 31.

Utz, H.F., Melchinger, A.E., 1996. PLABQTL: a program for composite interval ampping of QTL. Journal of Quantitative Trait Loci 2.

Valdes, A.M., Slatkin, M., Freimer, N.B., 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics 133, 737-749.

Valenzuela, N., 2000. Multiple paternity in side-neck turtles *Podocnemis expansa*: evidence from microsatellite DNA data. Molecular Ecology 9, 99-105.

Van Banning, P., 1985. Control of *Bonamia* in Dutch oyster culture. In: Ellis, A.E. (Ed.), Fish and Shellfish Pathology. Academic Press, London, pp. 393-396.

Van Banning, P., 1991. Observations on bonamiasis in the stock of the European flat oyster, *Ostrea edulis*, in the Netherlands, with special reference to the recent developments in Lake Grevelingen. Aquaculture 93, 205-211.

Van Ooijen, J.W., 2004. MapQTL 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma, B. V., Wageningen, Netherlands.

Van Ooijen, J.W., 2006. JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma, B. V., Wageningen, Netherlands.

Vercaemer, B., Spence, K.R., Herbinger, C.M., Lapègue, S., Kenchington, E.L., 2006. Genetic diversity of the European oyster (*Ostrea edulis* L.) in Nova Scotia: Comparison with other parts of Canada, Maine and Europe and implications for broodstock management. Journal of Shellfish Research 25, 543-551.

Verhaegen, D., Plomion, C., 1996. Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers. Genome 39, 1051-1061.

Viard, F., Bremond, P., Labbo, R., Justy, F., Delay, B., Jarne, P., 1996. Microsatellites and the genetics of highly selfing populations in the freshwater snail *Bulinus truncatus*. Genetics 142, 1237-1247.

Visscher, P.M., Haley, C.S., Knott, S.A., 1996. Mapping QTLs for binary traits in backcross and F2 populations. Genetical Research 68, 55-63.

Voorrips, R.E., 2002. MapChart: software for the graphical presentation of linkage maps and QTLs. The journal of heredity 93, 77-78.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M., 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23, 4407-4414.

Wada, K.T., 1984. Breeding study of the pearl oyster, *Pinctada fucata*. Bulletin of the National Research Institute of Aquaculture 6, 79-157.

Wada, K.T., 1986. Genetic selection for shell traits in the Japanese pearl oyster, *Pinctada fucata martensii*. Aquaculture 57, 171-176.

Waldbieser, G.C., Bosworth, B.G., Nonneman, D.J., Wolters, W.R., 2001. A microsatellite-based genetic linkage map for channel catfish, *Ictalurus punctatus*. Genetics 158, 727-734.

Wang, L., Song, L., Chang, Y., Xu, W., Ni, D., Guo, X., 2005. A preliminary genetic map of Zhikong scallop (*Chlamys farreri* Jones et Preston 1904). Aquaculture Research 36, 643-653.

Wang, S., Bao, Z., Pan, J., Zhang, L., Yao, B., Zhan, A., Bi, K., Zhang, Q., 2004. AFLP linkage map of an intraspecific cross in *Chlamys farreri*. Journal of Shellfish Research 23, 491-499.

Ward, R.D., English, L.J., McGoldrick, D.J., Nell, J.A., Thompson, P.A., 2000. Genetic improvement of the Pacific oyster *Crassostrea gigas* (Thunberg) in Australia. Aquaculture Research 31, 35-44.

Weber, J.L., 1990. Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. Genomics 7, 524-530.

Weber, J.L., May, P.E., 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. American Journal of Human Genetics 44, 388-396.

Weber, J.L., Wong, C., 1993. Mutation of human short tandem repeats. Human Molecular Genetics 2, 1123-1128.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G., Lathrop, M., 1992. A second-generation linkage map of the human genome. Nature 359, 794-801.

Whittaker, J.C., Harbord, R.M., Boxall, N., Mackay, I., Dawson, G., Sibly, R.M., 2003. Likelihood-based estimation of microsatellite mutation rates. Genetics 164, 781-787.

Wilkins, N.P., 1981. The rationale and relevance of genetics in aquaculture: An overview. Aquaculture 22, 209-228.

Williams, J.G., 1975. Sex and Evolution. Princeton University Press, Princeton, NJ.

Williams, R.J., 1970. Freezing tolerance in *Mytilus edulis*. Comparative Biochemistry and Physiology Part A 35, 145-161.

Wilson, J.H., Simons, J., 1985. Gametogenesis and breeding of *Ostrea edulis* on the west coast of Ireland. Aquaculture 46, 307-321.

Wilson, K., Li, Y., Whan, V., Lehnert, S., Byrne, K., Moore, S., Pongsomboon, S., Tassanakajon, A., Rosenberg, G., Ballment, E., Fayazi, Z., Swan, J., Kenway, M., Benzie, J., 2002. Genetic mapping of the black tiger shrimp *Penaeus monodon* with amplified fragment length polymorphism. Aquaculture 204, 297-309.

Winkelman, A.M., Peterson, R.G., 1994. Heritabilities, dominance variation, common environmental effects and genotype by environment interactions for weight and length in chinook salmon. Aquaculture 125, 17-30.

Wintero, A.K., Fredholm, M., Thomsen, P.D., 1992. Variable (dG-dT)n.(dC-dA)n sequences in the porcine genome. Genomics 12, 281-288.

Wong, A.K., Yee, H.A., van de Sande, J.H., Rattner, J.B., 1990. Distribution of CT-rich tracts is conserved in vertebrate chromosomes. Chromosoma 99, 344-351.

Woods, I.G., Kelly, P.D., Chu, F., Ngo-Hazelett, P., Yan, Y.L., Huang, H., Postlethwait, J.H., Talbot, W.S., 2000. A comparative map of the zebrafish genome. Genome Research 10, 1903-1914.

Yonge, C.M., 1960. Oysters. The New Naturalist, Collins, London, 209 pp.

Young, W.P., Wheeler, P.A., Coryell, V.H., Keim, P., Thorgaard, G.H., 1998. A detailed linkage map of rainbow trout produced using doubled haploids. Genetics 148, 839-850.

Yu, Z., Guo, X., 2003. Genetic linkage map of the eastern oyster *Crassostrea virginica* Gmelin. Biological Bulletin 204, 327-338.

Yu, Z., Guo, X., 2006. Identification and mapping of disease-resistance QTLs in the eastern oyster, *Crassostrea virginica* Gmelin. Aquaculture 254, 160-170.

Zane, L., Bargelloni, L., Patarnello, T., 2002. Strategies for microsatellite isolation: a review. Molecular Ecology 11, 1-16.

Zeng, Z.B., 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proceedings of the National Academy of Sciences of the USA 90, 10972-10976.

Zeng, Z.B., 1994. Precision mapping of quantitative trait loci. Genetics 136, 1457-1468.

Zhou, Z., Bao, Z., Dong, Y., Wang, S., He, C., Liu, W., Wang, L., Zhu, F., 2006. AFLP linkage map of sea urchin constructed using an interspecific cross between *Strongylocentrotus nudus* (♀) and *S. intermedius* (♂). Aquaculture 259, 56-65.

Zimmerman, A.M., Wheeler, P.A., Ristow, S.S., Thorgaard, G.H., 2005. Composite interval mapping reveals three QTL associated with pyloric caeca number in rainbow trout, *Oncorhynchus mykiss*. Aquaculture 247, 85-95.

Zouros, E., Pogson, G., 1994. The present status of the relationship between heterozygosity and heterosis. In: Beaumont, A.R. (Ed.), Genetics and Evolution of Aquatic Organisms. Chapman & Hall, London, pp.135–146.