

**Bangor University**

## **DOCTOR OF PHILOSOPHY**

### **Fisheries-induced evolution : a genetic approach using selection experiments on *Poecilia reticulata***

Van Wijk, Serinde

*Award date:*  
2011

*Awarding institution:*  
Bangor University

[Link to publication](#)

#### **General rights**

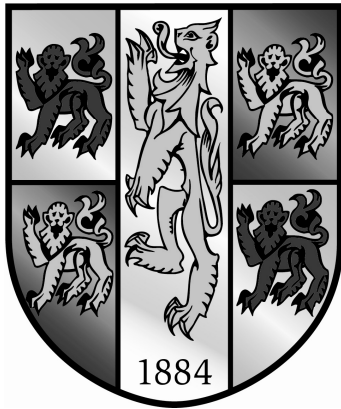
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

P R I F Y S G O L  
**BANGOR**  
U N I V E R S I T Y



# **Fisheries-induced evolution: a genetic approach using selection experiments on *Poecilia reticulata***

A thesis submitted to Bangor University for the degree of  
Doctor of Philosophy

By Serinde Joy van Wijk  
Oktober 2011  
Molecular Ecology and Fisheries Genetics Laboratory  
Bangor University  
School of Biological Sciences  
Environment Centre Wales  
BANGOR  
Gwynedd, LL57 2UW

## **DECLARATION AND CONSENT**

### **Details of the Work**

I hereby agree to deposit the following item in the digital repository maintained by Bangor University and/or in any other repository authorized for use by Bangor University.

**Author Name:** Serinde Joy van Wijk

**Title:** Fisheries-induced evolution: a genetic approach using selection experiments on *Poecilia reticulata*

**Supervisor/Department:** Prof. Gary R. Carvalho/ School of Biological Sciences

**Funding body:** Bangor University and the European Fisheries-Induced Evolution network

**Qualification/Degree obtained:** Doctor of Philosophy (PhD)

This item is a product of my own research endeavours and is covered by the agreement below in which the item is referred to as “the Work”. It is identical in content to that deposited in the Library, subject to point 4 below.

### **Non-exclusive Rights**

Rights granted to the digital repository through this agreement are entirely non-exclusive. I am free to publish the Work in its present version or future versions elsewhere.

I agree that Bangor University may electronically store, copy or translate the Work to any approved medium or format for the purpose of future preservation and accessibility. Bangor University is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

### **Bangor University Digital Repository**

I understand that work deposited in the digital repository will be accessible to a wide variety of people and institutions, including automated agents and search engines via the World Wide Web.

I understand that once the Work is deposited, the item and its metadata may be incorporated into public access catalogues or services, national databases of electronic theses and dissertations such as the British Library’s ETHOS or any service provided by the National Library of Wales.

I understand that the Work may be made available via the National Library of Wales Online Electronic Theses Service under the declared terms and conditions of use (<http://www.llgc.org.uk/index.php?id=4676>). I agree that as part of this service the National Library of Wales may electronically store, copy or convert the Work to any approved medium or format for the purpose of future preservation and accessibility. The National Library of Wales is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

**STATEMENT 1:**

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless as agreed by the University for approved dual awards.

Signed ..... (candidate)

Date .....

**STATEMENT 2:**

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).

All other sources are acknowledged by footnotes and/or a bibliography.

Signed ..... (candidate)

Date .....

**STATEMENT 3:**

I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loans and for electronic repositories **after expiry of a bar on access**.

Signed ..... (candidate)

Date .....

**STATEMENT 4:**

I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University when the approved **bar on access** has been lifted.

**IN ADDITION TO THE ABOVE I ALSO AGREE TO THE FOLLOWING:**

1. That I am the author or have the authority of the author(s) to make this agreement and do hereby give Bangor University the right to make available the Work in the way described above.
2. That the electronic copy of the Work deposited in the digital repository and covered by this agreement, is identical in content to the paper copy of the Work deposited in the Bangor University Library, subject to point 4 below.
3. That I have exercised reasonable care to ensure that the Work is original and, to the best of my knowledge, does not breach any laws – including those relating to defamation, libel and copyright.
4. That I have, in instances where the intellectual property of other authors or copyright holders is included in the Work, and where appropriate, gained explicit permission for the inclusion of that material in the Work, and in the electronic form of the Work as accessed through the open access digital repository, *or* that I have identified and removed that material for which adequate and appropriate permission has not been obtained and which will be inaccessible via the digital repository.
5. That Bangor University does not hold any obligation to take legal action on behalf of the Depositor, or other rights holders, in the event of a breach of intellectual property rights, or any other right, in the material deposited.
6. That I will indemnify and keep indemnified Bangor University and the National Library of Wales from and against any loss, liability, claim or damage, including without limitation any related legal fees and court costs (on a full indemnity bases), related to any breach by myself of any term of this agreement.

Signature: ..... Date : .....

*How fleeting are the wishes and efforts of man! How short his time!  
and consequently how poor will his products be, compared with those  
accumulated by Nature*

Charles Darwin, 1859

To my brother, who taught me to appreciate every step on the way

## **ABSTRACT**

Increasing evidence suggests that size-selective mortality imposed by commercial fishing results in directional changes in life histories of exploited species, including effects on maturation age, growth rate and body size. Whether these changes are the result of fisheries-induced evolution or other selective pressures in the natural environment and/or phenotypic plasticity continues to be the subject of much debate. Currently, molecular genetic data revealing fisheries-induced shifts at candidate loci are lacking.

Here, the hypothesis that harvesting can induce genetic change over few generations was tested directly by subjecting laboratory-reared offspring of wild-caught Trinidadian guppies, *Poecilia reticulata*, to divergent, size-specific selection over three generations. The smallest/ largest/ random twenty percent of males was selected each generation and changes in standard length, as well as in the frequencies of alleles at neutral microsatellite loci and putative candidate genes for selection were recorded.

Significant divergence between differently selected lines was observed for male standard length ( $\pm 7\%$ ), size ( $\pm 8-12\%$ ) and age ( $\pm 4-6\%$ ) at maturation, compared to only 1% change in standard length over generations in the random breeding control line. Significant drift between lines, but no genetic erosion or inbreeding, was apparent over generations at microsatellite markers. Signatures of selection and significant genetic divergence between selected lines were detected at five out of 17 putative candidate loci (Pr39, M9, M30, M987 and prolactin) which confirmed strong Y-linkage of genes underlying male body size in guppies, as indicated by the phenotypic data. Additionally, significant genotype-phenotype associations were obtained for twelve of the candidate genes. For two of these loci (M30 and M1046) an association between the same single nucleotide polymorphism and a QTL for standard length had been observed previously.

To our knowledge, this is the first study where selection on body size with known intensity and direction has been compared directly with both a phenotypic response and changes at individual genetic marker loci. Hereby, this study forms one of the first pieces of molecular genetic evidence for fisheries-induced evolution: by demonstrating that phenotypic shifts in body size resulting from size-selective harvesting, comparable to commercial fisheries in principle, are underlain by quantifiable genetic change.



## **ACKNOWLEDGEMENTS**

The work presented in this thesis would not have been completed with the same level of satisfaction as it has today, were it not for the contributions of some extraordinary people, whom I would like to thank on these pages.

### **THE MINDS**

In the first place, I am grateful to my supervisors: Gary Carvalho, Si Creer, Cock van Oosterhout and Martin Taylor, for their non-stop confidence in my abilities and continuous support throughout the past four years. I feel incredibly lucky to have had such a diverse team of supervisors, having the whole range of expertise from theory to experimental design and analysis between the four of you. In particular, I would like to thank you for having the long-term vision assuring this work would succeed, thereby keeping me going through times when I could not even imagine it to. Your guidance and equally strong capacity to leave me to learn at my own pace, have been inspirational.

Years ago, it was Franjo Weissing, whose keen enthusiasm is so very contagious, who encouraged me to pursue a PhD. It was during my stay in his “TheoBio” group in Groningen, still a very dear memory, that I discovered I enjoyed research and actually was good at it. I would like to thank you for our many discussions, all your support back then and now, and most of all for your gentle encouragements for me to have confidence in my talents. It is still growing.

The Molecular Ecology and Fisheries Genetics lab hosts a fantastic group of people, who are always willing to share their expertise and patiently explain whatever needs explaining. I am particularly thankful to Sonia Pascoal, Mrinalini and Markos Alexandrou for advice on pretty much anything involving analysis, lab work or writing a thesis, and to Delphine Lallias and Nik Dawney for making Genemapper doable. Special thanks go to Niklas Tysklind for his microsatellite expertise and for pointing my attention towards allele 174 at locus Pr39, helping me understand why it was not annoyingly misbehaving but actually very interesting.

I would like to thank Hazel Perry for building further on my project and continuing the work with the selection lines. It is very satisfactory to see my efforts will be of use to someone else and many of the things I have not been able to investigate will still be explored.

This research was funded by Bangor University and the European Fisheries-induced Evolution network.

## THE HANDS

I am indebted with the many people who put their time and effort into helping me out with so many of the practical aspects of this PhD.

In the first place, my sincere gratitude goes out to Indar Ramnarine and Rajindra Mahabir, from the University of the West Indies in St. Augustine, Trinidad, without whom I would not have had any guppies to work with in the first place, and who made a perspective-widening field trip to Trinidad possible. Thank you as well Gary, John, Martin and Marj, with little Finn on board, for joining me on the expedition.

Christine Dreyer and Margarete Hoffmann, from the Max Planck Institute in Tübingen, Germany, provided me with candidate markers and primers from their own work, which have been invaluable in the molecular aspects of my project and allowed for some of the major findings of this study. I hope my work will be of use to you, and there will be further collaborations in the future.

I am grateful to Emlyn Roberts, Wynn and Katie Johnston, Ann Pennell, Ashley Tweedale, John Evans and Alan Jones, the hidden faces behind anything practical happening within SBS, for all your countless solutions to the inevitable – but nevertheless unexpected – little disasters I managed to create and you managed to deal with so well.

My dear aquarium buddies; Dawn, Gavan, Alix, Markos and Tara, who made the time in the Brambell basement so much more enjoyable; with whom I could share thoughts, enthusiasm and frustrations in long hours without real daylight; who fed and looked after my fish on so many occasions and who made the whole experience of working with live animals so much more enjoyable. Thank you, I would not have lasted otherwise.

The husbandry and measuring of fish has been done for a great deal by two placement students and by Hwyl Lloyd, who put in many hours and whose feedback helped me to get on top of what I was doing, and what needed to be done.

Very special thanks go out to Wendy Grail and Jan Albin, the two backbones of MEFGL who, besides the myriad of help and support they have provided doing their own jobs, both have spent hours and hours netting fish, photographing them and assisting me with some of the most tedious aspects of my experiments.

Above all, this project would not have been completed anywhere near its current form without the help of Marta Rodrigues, who just appeared one day; keen to learn, keen to help, igniting my own enthusiasm again. Thank you so much. You have done more for my PhD than I thought possible, the best thing being we both enjoyed it so much.

## THE HEARTS

All the work being done, in the end it only matters if it can be enjoyed, shared, learned from, and enriches your life beyond the boundaries of what you started. That process involves the best of people all around you and makes it all worthwhile.

It involved the good company of all those mentioned here before and especially my dear friends: Joana, Muns, Markos, Niklas, Sonia, Sophie, Miranda, Dawn, Greg, Gavan and Alix, Iliana, Sarah and Wol, the Capoeira crew, Bryn and T, Jacqui, Hana, Eilir, Jenny and Brew, Lucy, Helen, Andy, Hillary, Mark and all the dear and fantastic members of the Menai Massive. Thank you Rhod, for enriching my life in so many ways. Thanks to my fellow rock addicts: Emlyn, for teaching me and bearing with my slightly bad judgement when I get excited about a route; Matt, for still my best ever climbing experience; Simon, Sophie, Si, Rob, Hillary and Beattie, Marija, Johnny, Liam and Toby for countless adventures. How I am looking forward to meet you all more now this work is done!

A special commendation goes to the Outstanding Support provided by Polly and Martin, Si, and Emlyn, on the separate occasions when I needed lifts from various climbing venues to A&E. And to Gjalt, for housing me (in dire need of refuge at the time) in his office for a whole month.

Even more I want to thank those old friends who have never moved one step further away, no matter how great the distance. My three other quarters; Lempie, Linde and Marlieschen – we all change, but We doesn't change. My two fantastic brother-in-laws Gjalt and Leo, Machteld, Ingrid, Thor, Solief, Niek, Renske, Roos, Chantallala, Menno, all the superb RKZers; thank you so much for your friendship and the immortal times, of which I hope many more are still to come.

More than anything I thank my parents, my brother and sisters and my grandmother, for their support, understanding, encouragement and the unconditional bond that we share. It is fulfilling to experience it is only getting stronger, no matter where we are.

## **TABLE OF CONTENTS**

<b>List of Tables</b>	<b>xiv</b>
<b>List of Figures</b>	<b>xv</b>
<b>Glossary</b>	<b>xvi</b>
<b>Chapter 1. General Introduction</b>	<b>1-35</b>
<b>1.1 Life history evolution and natural selection in the wild</b>	<b>2-3</b>
1.1.1 Adaptation to changing environments	2
1.1.2 Understanding adaptation	3
<b>1.2 Contemporary evolution</b>	<b>4-18</b>
<u>Box 1</u> . Strategies for identifying the genetic basis to adaptation and fitness-related traits	6
1.2.1 Harvesting of marine resources	7
1.2.2 Fishing as a selective force	9
1.2.3 Fisheries-induced change in life history traits	10
1.2.4 Fisheries-induced evolution?	11
1.2.5 Discriminating between evolutionary change and phenotypic plasticity	13
1.2.6 The use of probabilistic reaction norms for inferring genetic change	15
1.2.7 The missing link: conclusive evidence of genetic change	17
<b>1.3 Selection experiments and the study of adaptation</b>	<b>19-21</b>
1.3.1 Advantages and limitations to the use of selection experiments	19
1.3.2 Experimental studies on fisheries-induced evolution	20
<b>1.4 The Trinidadian guppy, <i>Poecilia reticulata</i>, as a model species</b>	<b>22-31</b>
1.4.1 <i>Poecilia reticulata</i> as a research model	22
1.4.2 Population distribution; phylogenetic history and connectivity	24
1.4.3 Predation and life history strategies	25
1.4.4 Transplant experiments	28
1.4.5 Environmental differences between upland and lowland habitats	28
1.4.6 Predator introduction experiment	29
1.4.7 <i>Poecilia reticulata</i> as a genetic and genomic model	30
<b>1.5 Aims of the study and outline of experiments</b>	<b>32-34</b>
1.5.1 Aims and objectives	32
1.5.2 Underlying rationale	32
1.5.3 Outline of the experiments and thesis	35

<b>Chapter 2. Phenotypic divergence of male <i>P. reticulata</i> in response to three generations of artificial selection on male standard length</b>	<b>36-73</b>
<b>2.1 Introduction</b>	<b>37-42</b>
2.1.1 Phenotypic change and micro-evolution in natural populations	37
2.1.2 The required evidence for quantifying fisheries-induced evolution	38
2.1.3 Artificial selection experiments for studying fisheries-induced evolution	40
<u>Box 2.1.</u> Terminology in quantitative genetics	42
<b>2.2 Experimental design</b>	<b>43-45</b>
2.2.1 Rationale behind chosen selection regime	43
2.2.2 Minimising effects of drift and sexual selection	43
2.2.3 Minimising maternal effects and inbreeding	44
<u>Box 2.2.</u> Maternal effects	45
<b>2.3 Materials and methods</b>	<b>46-52</b>
2.3.1 Sampling and transport of fish to the Bangor University aquarium facility	45
2.3.2 Health monitoring	46
2.3.3 Rearing protocol	46
2.3.4 Rearing of random breeding generations: removal of maternal effects and increasing population size in generations F1-F2	47
2.3.5 Establishing selection lines in generation F3	47
2.3.6 Maintenance of selection lines in generations F4-F6	48
2.3.7 Estimation of maturation age and size in the F2 and F6 generation	49
2.3.8 Measurements	50
2.3.9 Statistical analyses	52
<b>2.4 Results</b>	<b>53-64</b>
2.4.1 Breeding efficacy	53
2.4.2 Experimental results – random breeding generations	54
2.4.3 Experimental results – response to selection	55
2.4.4 Estimates of heritability and quantitative genetic parameters	60
2.4.5 Rate of evolution	63
<b>2.5 Conclusions and Discussion</b>	<b>65-74</b>
2.5.1 Response to selection	65
2.5.2 Heritability of standard length	67
2.5.3 Variance components and their response to selection	69
2.5.4 Evolutionary rate of standard length	71
2.5.5 Overall conclusions	73

<b>Chapter 3. Genetic dynamics of selection lines of <i>P. reticulata</i> I – Genetic variation at neutral microsatellite loci and drift between selected lines</b>	<b>75-103</b>
<b>3.1 Introduction</b>	<b>76-78</b>
3.1.1 Microsatellites for studying genetic diversity of exploited fish species	76
3.1.2 Microsatellites for studying genetic diversity in artificial selection experiments	77
<b>3.2 Materials and methods</b>	<b>79-83</b>
3.2.1 Origin of fish and rearing protocol	79
3.2.2 DNA extraction	79
3.2.3 Microsatellite marker choice and PCR protocols	79
3.2.4 Genetic analyses	80
3.2.5 Paternity analyses	83
<b>3.3 Results</b>	<b>84-95</b>
3.3.1 Loci amplification	84
3.3.2 Null alleles, Hardy-Weinberg equilibrium and linkage	84
3.3.3 Genetic diversity	86
3.3.4 Genetic structure	90
3.3.5 Paternity reconstruction	93
<b>3.4 Conclusions and discussion</b>	<b>96-103</b>
3.4.1 The effect of selection on neutral microsatellite loci	96
3.4.2 The magnitude of drift through captive breeding	97
3.4.3 Paternity reconstruction	98
3.4.4 Heritability of standard length	100
3.4.5 Overall Conclusions	102
<b>Chapter 4. Genetic dynamics of selection lines of <i>P. reticulata</i> II – Identification of a response to selection at candidate loci</b>	<b>104-159</b>
<b>4.1 Introduction</b>	<b>105-110</b>
4.1.1 A candidate gene approach to identifying functional genetic variation	105
<u>Box 4.</u> Candidate genes for growth and reproduction	107
4.1.2 A context for the identification of polymorphisms and genetic association studies	109
4.1.3 A candidate gene approach to investigating the response to selection on body size	109
<b>4.2 Materials and methods</b>	<b>111-118</b>
4.2.1 DNA sampling and extraction	111
4.2.2 Candidate marker choice	111
4.2.3 Sequencing reactions and PCR protocols	112
4.2.4 SNP discovery	113
4.2.5 Genetic diversity and Linkage disequilibrium	114
4.2.6 Rationale underlying candidate marker analysis	114
<b>4.3 Results</b>	<b>119-143</b>
4.3.1 SNP discovery and genotyping success rate	119
4.3.2 Genetic diversity and linkage disequilibrium	121
4.3.3 Hardy-Weinberg disequilibrium and $G_{IS}$	123
4.3.4 Genotypic differentiation	126

4.3.5	AMOVA	131
4.3.6	Outlier analyses: multiple-marker based neutrality tests	132
4.3.7	Association analysis	140
<b>4.4</b>	<b>Conclusions and discussion</b>	<b>144-159</b>
4.4.1	Evidence for selection at candidate loci	144
4.4.2	Congruence and discrepancies between analyses	148
4.4.3	The genetic basis of male body size: differences between small and large size	154
4.4.4	The genetic basis of male body size: Y-linkage of candidates	156
4.4.5	Putative gene function of major candidate genes	157
4.4.6	Overall conclusions	158
<b>Chapter 5. General Discussion</b>		<b>160-170</b>
<b>5.1</b>	<b>Summary of main results and novelties of the research</b>	<b>161-162</b>
<b>5.2</b>	<b>Broader relevance of the obtained results</b>	<b>163-166</b>
5.2.1	Life history evolution of Trinidadian guppies	163
5.2.2	Fisheries-induced evolution	164
<b>5.3</b>	<b>Future directions</b>	<b>167-170</b>
5.3.1	Characterisation of genetic differences between selection lines	167
5.3.2	Other putative traits under selection	168
5.3.3	The significance of evolutionary principles for fisheries management	169
<b>References</b>		<b>171-197</b>
<b>Appendices</b>		<b>198-244</b>
Appendix I.	Primers and sequences	199
Appendix II.	Microsatellite allele frequencies, $H_O$ and $H_S$ per locus	202
Appendix III.	Full AMOVA Tables	207
Appendix IV.	Pairwise $F_{ST}$ and $G''_{ST}$ per locus	213
Appendix V.	Haplotypes for five Growth Hormone SNPs	216
Appendix VI.	Full tables of genotypic differentiation	217
Appendix VII.	Fdist outlier analyses figures	225
Appendix VIII.	Raw and normalised $\ln(RH)$ data	232
Appendix IX.	Association analyses models of inheritance	235

## **LIST OF TABLES**

Table 2.1	Male standard length in generations F3-F6	56
Table 2.2	Summary of linear regression analysis of male st. length in the F3-F6	55
Table 2.3	Summary of linear regression analysis of female st. length in the F3-F6	55
Table 2.4	Heritability estimates from the linear regression of R/S	61
Table 2.5	Evolutionary rate of standard length in darwins and haldanes	64
Table 3.1	Overview of population genetic parameters of eight microsatellite loci for each selection line in each generation (F0-F6)	88
Table 3.2	Number of alleles per locus, line and generation a eight microsatellite loci	90
Table 3.3	$F_{CT}$ , $F_{ST}$ and $F_{SC}$ for a hierarchical AMOVA of microsatellite variation per generation in the F3-F6 and for a temporal AMOVA	91
Table 3.4	Summary of paternity assignments of male fish	95
Table 4.1	Overview of candidate genes investigated for a response to selection	110
Table 4.2	Allele frequencies at candidate loci in generations F0, F2, F5 and F6	120
Table 4.3	Candidate markers showing significant linkage disequilibrium	121
Table 4.4	Observed and expected heterozygosity for candidate loci	130
Table 4.5	Significant values of $G_{IS}$ per locus and generation with different levels of statistical support	125
Table 4.6	Summary of candidate loci showing significant genotypic differentiation between population pairs	127
Table 4.7	$F_{CT}$ , $F_{ST}$ and $F_{SC}$ per locus for hierarchical AMOVA of candidate gene variation in the F5 and F6	133
Table 4.8	Overview of significant outlier loci between population pairs identified by LOSITAN	135
Table 4.9	Overview of significant outlier loci between population pairs identified by Schlötterer's method	139
Table 4.10	Summary of candidate loci identification by different analyses	144



## **LIST OF FIGURES**

Fig. 1.1	Citation report for studies on fisheries(-)induced evolution	14
Fig. 1.2	Representation of a probabilistic maturation reaction norm	15
Fig. 1.3	Worldwide distribution of <i>Poecilia reticulata</i>	22
Fig. 1.4	Comparison between commercial harvesting mortality and differential mortality on guppies in high- and low predation habitats	27
Fig. 1.5	Comparison between artificial and natural harvesting selection regimes	34
Fig. 2.1	Representation of the knife-edge selection procedure used	48
Fig. 2.2	Schematic representation of the experimental design	51
Fig. 2.3	Male maturation and gonopodium development of <i>P. reticulata</i>	50
Fig. 2.4	Overview of male standard length in random breeding generations	53
Fig. 2.5	Male growth proportional to maturity status in generation F2	54
Fig. 2.6	Male standard length in different selection lines in generations F3-F6	56
Fig. 2.7	Female standard length in different selection lines in generations F3-F6	58
Fig. 2.8	Male size and age of maturation in generation F6	60
Fig. 2.9	Response (R) to selection as a function of the selection differential (S)	61
Fig. 2.10	Changes in $V_P$ and $V_A$ in selection lines over generations F3-F6	62
Fig. 3.1	$H_O$ , $H_S$ and $G_{IS}$ over eight microsatellite loci for all lines and generations	87
Fig. 3.2	Allelic Richness over eight microsatellite loci for all lines and generations	89
Fig. 3.3	Hierarchical AMOVA of microsatellite variation per generation in the F3-F6 and for a temporal AMOVA	91
Fig. 3.4	Pairwise $F_{ST}$ and $G''_{ST}$ per generation between selection lines over eight microsatellite loci	92
Fig. 3.5	Father-son regression of standard length	95
Fig. 4.1	Minor allele frequencies over generations at 17 candidate loci	128
Fig. 4.2	Observed and expected homozygosity of allele "174" at locus Pr39	130
Fig. 4.3	Hierarchical AMOVA of candidate gene variation	132
Fig. 4.4	Pairwise $F_{ST}$ and $G''_{ST}$ per generation between selection lines over all microsatellite and all candidate loci	134
Fig. 4.5	Pairwise $G''_{ST}$ per locus for all microsatellite and all candidate loci	134
Fig. 4.6	Normalised distribution of $\ln(RH)$ for all individual loci	137
Fig. 4.7	Significant genotype-phenotype associations for candidate loci	142
Fig. 4.8	Correlation of allele 174 frequency at Pr39 and male standard length	148

## GLOSSARY

<b>AMOVA</b>	Analysis of molecular variance
<b>ANOVA</b>	Analysis of variance
<b>AT</b>	Annealing temperature
<b>BLAST</b>	Basic Logical Alignment Search Tool
<b>bp</b>	Basepair
<b>cDNA</b>	Complementary DNA
<b>cds</b>	Coding sequence
<b>CTAB</b>	Hexadecyltrimethylammomium Bromide
<b>d</b>	Darwins
<b>DNA</b>	Deoxyribonucleic acid
<b>dNTP</b>	Deoxynucleoside triphosphate
<b>FAO</b>	Food and Agriculture Organization
<b>FBPase</b>	fructose-1,6-bisphosphatase 1
<b>FIE</b>	Fisheries-induced evolution
<b>GH</b>	Growth hormone
<b><math>h^2</math></b>	Heritability
<b><math>H_o</math></b>	Observed heterozygosity
<b><math>H_s</math></b>	Expected heterozygosity
<b>HW</b>	Hardy Weinberg
<b><math>i</math></b>	Selection intensity
<b>IGF</b>	Insulin growth factor
<b>Indel</b>	Insertion-deletion
<b>LD</b>	Linkage disequilibrium
<b>MC4R</b>	Melanocortin-4 receptor
<b>MgCl<sub>2</sub></b>	Magnesium Chloride
<b>NGS</b>	Next generation sequencing
<b>ORF</b>	Open reading frame
<b>PCR</b>	Polymerase chain reaction
<b>PMRN</b>	Probabalistic maturation reaction norm
<b>QTL</b>	Quantitative trait loci
<b><math>R</math></b>	Selection response
<b>RAD</b>	Restriction site-associated DNA
<b>RIPL1</b>	rab interacting lysosomal-like protein-1
<b><math>S</math></b>	Selection differential
<b>S.E.</b>	Standard error of the mean
<b>SL</b>	Standard length
<b>SNP</b>	Single nucleotide polymorphism
<b>St. dev.</b>	Standard deviation
<b><math>V_A</math></b>	Additive genetic variance
<b><math>V_P</math></b>	Phenotypic variance

# **1. GENERAL INTRODUCTION**

## **1.1 LIFE HISTORY EVOLUTION AND NATURAL SELECTION IN THE WILD**

### **1.1.1 Adaptation to changing environments**

Individuals adapt differently to different environments. Understanding the rate and mechanisms by which populations adapt to changing environments touches on the very basis of evolutionary biology and has been the focus of frontline research for decades. It was Darwin's personal observations of local adaptation of beak shape of Galapagos finches on different islands, which led directly to the theory of evolutionary change by natural selection (Darwin 1859). Since this early work, a vast amount of understanding of fundamental evolutionary processes and the dynamic interactions of populations and their environment has been achieved by studying local adaptation and life history evolution in the wild (Endler 1986; Hendry et al. 2011).

What is adaptation? It is important to have a good understanding and precise definition of this deceptively straightforward term. Throughout the present study, adaptation is regarded as a process that increases fitness in any given environment (modified from Hendry et al. 2011). Using this basic definition it becomes instantly clear that traits closely related to fitness, such as life history traits, should be strong targets of selection and thus, changes in life history traits will often be involved in adaptation. It is also important to emphasise that this definition includes both genetic and non-genetic adaptations, implying that not only evolution in a strict, genetic, sense but also phenotypic plasticity (Stearns 1989; Ghalambor et al. 2007) can underlie adaptation. In contrast, throughout this thesis, only genetic change is meant when discussing evolutionary change.

At a phenotypic level, adaptation through genetic change and phenotypic plasticity may have similar outcomes and can be hard to distinguish, for example when countergradient variation is present (Conover and Present 1990; Conover and Schultz 1995) or when environmental change over the study period results in cryptic evolution (Merilä et al. 2001a). From an evolutionary perspective however, the difference between genetic and plastic change is crucial and it can be appreciated

only by an understanding of the mechanisms underlying adaptations (Vasemägi and Primmer 2005; Latta et al. 2007; Gienapp et al. 2008). Understanding the (genetic) basis of adaptations can yield insights into speciation processes, the heritability of adaptations, (that is; whether they will be passed on to future generations), if adaptations are reversible and whether adapted individuals and populations are likely to be able to cope with further changing environments.

### **1.1.2 Understanding adaptation**

To appreciate the mechanisms involved in local adaptation, an understanding is needed of the ecological context of-, or differences between-, contrasting habitats and localities of interest. Furthermore, the genetic differences between, and colonisation history of, the study populations should be assessed. It is the combination of these factors, which often interact in a complex fashion, that shapes population differentiation and determines the rate and nature of adaptation (Shaw et al. 1994; Li & Merilä 2001; Merilä et al. 2001b; Hilborn et al. 2003; Garant et al. 2005).

Molecular techniques are indispensable tools to uncover and understand adaptation processes and to predict how and when populations will adapt to changing environments (Garant and Kruuk 2005; Hendry et al. 2011; Le Rouzic et al. 2011). Neutral molecular genetic markers, such as microsatellites, have provided information on population structure and differentiation in a vast range of species, from skin colour in humans (Jablonski 2004), to local adaptation in seemingly unstructured habitats like the sea (Hansen et al. 2001; Nielsen et al. 2009a).

Recently, technological advances at a quantitative and molecular level have further enhanced our understanding of evolutionary change and the genetic basis of adaptation (Box 1). Until recently, by far the majority of such work has been performed on model organisms, where a wealth of genetic resources and the availability of inbred lines and long-standing selection experiments facilitate powerful associations and genetic scans (Kadarmideen and Janss 2007; Burke et al. 2010). However, next generation sequencing technologies and advanced molecular

techniques are aiding the discovery of genes and genomic regions under selection in a vast range of species (Reusch and Wood 2007; Stapley et al. 2010). Thus, the ultimate goal of characterising the nature and diversity of adaptive genetic variation and the genetic basis of fitness differences between individuals in different environments, is becoming feasible for an increasing number of species (Luikart et al. 2003; Ellegren and Sheldon 2008; Stinchcombe and Hoekstra 2007; Tenaillon and Tiffin 2008). A renewed interest into functional genetic differences between populations has arisen as we can now finally begin to appreciate truly how traits evolve and what genetic correlations and constraints shape the evolution of (life history) traits (Vasemägi & Primmer 2005; Slate et al. 2008; Nielsen et al. 2009a).

For example, different modes of selection at polymorphic sites in the rhodopsin gene have been demonstrated to act in gobies (*Pomatoschistus minutes*) inhabiting water bodies that differ in turbidity and vision (Larmuseau et al. 2009); differences in gene expression profiles explain colonisation patterns in butterflies (Wheat et al. 2011); genome scans have allowed for detecting candidate loci involved in freshwater adaptation in stickleback (Hohenlohe et al. 2010) and horn morphology in wild Soay sheep (Johnston et al. 2011); they are but a few of the breadth of excellent studies in recent years that have made use of molecular tools to answer fundamental questions in evolutionary biology.

## **1.2 CONTEMPORARY EVOLUTION**

Whilst initially regarded as a slow process taking many millions of years or hundreds of generations (Darwin 1859; Wells et al. 1934; Mayr 1966), the realisation has risen that evolution can happen quickly in response to rapidly changing environments (Hendry and Kinnison 2001; Reznick and Ghalambor 2001; Reznick and Ghalambor 2005b). Numerous examples exist that demonstrate adaptation can be a rapid process (Endler 1980; Reznick et al. 1990; Avilés et al. 2006; Palkovacs et al. 2008; Barrett et al. 2010). This understanding has been further fuelled by investigations of how the natural world is affected by human activities and disturbances. Adaptive responses to, for example, pollution (Kettlewell 1955, 1973) fisheries (Sharpe and Hendry 2009), hunting (Coltman et al. 2003); insect control (McKenzie and

Batterham 1994; Raymond et al. 2001) and pesticide resistance (Heap 1997) have demonstrated unequivocally that human activities affect habitats worldwide at a pace rarely observed in history (Palumbi 2001; Hendry et al. 2008).

In particular, human harvesting is a selective force far more direct and consistent than many natural selective drivers (Coltman et al. 2003; Allendorf et al. 2008; Coltman 2008; Darimont et al. 2009) and, in order to safeguard biodiversity and yield, there is an urgent need to understand how and at what rate human disturbances affect ecosystem stability and resilience (Scheffer et al. 2001), species distribution, (Møller and Mousseau 2008, 2009) population dynamics, trophic interactions (Estes et al. 1998; Lande 1998) and evolution of traits (Stockwell et al. 2003).

## **BOX 1. STRATEGIES FOR IDENTIFYING THE GENETIC BASIS TO ADAPTATION AND FITNESS-RELATED TRAITS**

Recently, major technological advances have been made aiding the identification of genetic variation and loci underlying functionally important traits in non-model organisms. Such endeavours are important, as they contribute to our understanding of fundamental evolutionary processes and allow us to predict the long-term consequences of human disturbances. In particular, the possibility to create and access vast amounts of genetic and genomic resources has improved our understanding of these processes. Here, these tools are briefly summarised.

Broadly, approaches for identifying functionally important genetic variation can be classified in two main groups. Firstly, bottom up approaches do not require any *a priori* knowledge on the function and properties of genes of interest and use genome-wide detection techniques (Storz 2005; Butlin 2010) that create and screen large amounts of random markers for  $F_{ST}$  outliers in individuals from contrasting environments (e.g. Bonin et al. 2006). This approach yields information on genomic regions putatively under selection, which can then be investigated in more detail to identify loci, and ultimately genes likely to be involved in adaptation (e.g. Wood et al. 2008). Alternatively, when pedigrees are available, associations with phenotypic traits of interest on a high-density linkage map can be made directly (e.g. Rogers and Bernatchez 2005; Johnston et al. 2011).

Secondly, when prior knowledge convincingly suggests that specific loci are likely to be involved in the trait(s) of interest, a top down approach may be preferred, in which the genes of interest are investigated directly for polymorphisms and, if found, followed by exploration of the associations of polymorphisms with relevant phenotypes (Piertney and Webster 2010). When sequence information on the target gene is available for the species of interest, obtaining primer information can be straightforward (e.g. Hemmer-Hansen et al. 2011; this study) but alternatively, published sequences from related species can be used to design primers in conserved regions that cross-amplify in the study species (Primmer et al. 2002; Aitken et al. 2004).

Both approaches have profited immensely from recent advances in sequencing technology allowing for large numbers of markers to be identified and/ or genotyped simultaneously (Mardis 2008; Stapley et al. 2010). Cost-effectively, these approaches are more worthwhile in studies where few individuals and large numbers of markers per individual are required, such as mapping studies or cases where pooled population samples can be used (Burke et al. 2010). When investigating allele frequencies for numerous samples, next generation sequencing still is very costly. A promising new approach in this context is restriction-site associated DNA (RAD) tagging, where numerous SNPs are detected and sequenced simultaneously for large samples in a single sequencing run (Baird et al. 2008). Especially when a reference genome is available, SNP discovery can be realised with relatively light sequencing effort. Additionally, the approach can be used to confirm the role of candidate loci from mapping studies (Holenhohe et al. 2010). Even without the availability of a sequenced genome, the approach is still very powerful in locating functional regions of differentiation from other genomic sources and/ or crosses (Baird et al. 2009; Rowe et al. 2011).

The applications of next generation sequencing (NGS) in recent years predict a surge in knowledge on the genetic basis of fitness variation in the wild. With third generation sequencing close behind it, providing longer read lengths and increased flexibility in *de novo* assembly of genomic data (Stapley et al. 2010), adaptation genomics are expected to become accessible for a rapidly increasing number of natural populations.



### **1.2.1 Harvesting of marine resources**

With respect to yield, fisheries occupy a unique position between other harvested resources, as fish are the only commercial food source that is harvested in the wild on a large scale (as opposed to small scale harvesting such as wild fungi). The tremendous catches by commercial fisheries embody the single largest mortality factor on marine fish post larval stages. Global catches peaked in 1996 with 86.3 million tons of fish. 79.5 Million tons were caught in 2008 and almost 50% of the world's population consumes 15% of their animal protein by eating fish (FAO 2010). In the last century, fisheries have resulted in overexploitation over 30% of commercially valuable fish species with another 50% being exploited up to the limits of their sustainable maxima (FAO 2010).

Fisheries are the single largest mortality factor on adults of harvested stocks, and as such, fisheries form one of the strongest selective forces on wild fish populations observed in the last 100 years (Fenberg and Roy 2008). An increasing body of evidence shows that selection by fishing has negative impacts not only on yield and population sizes (Heino 1998; Law and Grey 1989; Conover and Munch 2002; Ernande et al. 2004) but also life history traits (Jorgensen et al. 2007) and there is an urgent need to understand what adaptations are taking place in harvested species in response to fisheries, both at a phenotypic and a genetic level. In addition to affecting target species, fisheries affect their communities (Jennings et al. 1999; Shackell et al. 2009), resulting in additional loss of recovery potential (Hutchings 2000), resilience (Hsieh et al. 2006) and ecosystem productivity (Worm et al. 2006). Collapses of fish stocks (Beverton 1990; Myers et al. 1997; Hutchings 2000; Dulvy et al. 2003; FAO 2010; Pinsky et al. 2011) combined with globally declining catches (Myers and Worm 2003; Pauly et al. 2003; FAO 2010) show that our understanding of the impact of fisheries on marine resources has been insufficient for their descent management.

The reason for this failure to understand – and predict – the dynamics of marine resources lies partly in the short-sighted vision of some fisheries managers that seem

incapable of safeguarding yield for themselves or future generations<sup>1</sup> (Rutter 1902; Hardin 1968; Botsford 1997; Beverton 1998). Secondly, its explanation lies in the difficulties involved in studying dynamics of underwater populations with very large census sizes, high juvenile mortality and large inter-annual variation in recruitment, survival and population sizes (Winemiller and Rose 1992; Hedgecock 1994; Iles and Beverton 2000; Berkeley et al. 2004). Life history strategies of many harvested marine fish are iteroparous, so-called “bet-hedging” strategies: where reproductive success is highly variable and spread out over a long reproductive period following maturation late in life (Fogarty et al. 1991; Winemiller and Rose 1992). Furthermore, it is often the older and larger individuals that produce the most viable offspring (Trippel 1995; Walsh et al. 2006; Johnson et al. 2011). For the majority of harvested fish species however, detailed information on reproductive behaviour, stock identity and connectivity, migration, compensatory and depensatory processes and trophic interactions is rare at best (Conover et al. 2006; Liermann and Hilborn 2001; Rose et al. 2001). However, all these factors can greatly influence recruitment and survival; violation of assumptions on any on these processes will affect the reliability of predictions of current and future stock sizes (Fogarty et al. 1991; Botsford 1997; Rijnsdorp et al. 2007). Although considerable effort is being made to infer historic population sizes from historical data (Jackson 2001; Lotze and Worm 2009) such information is seldom available – nor used as a baseline reference, even if accessible (Pauly 1995). As research surveys typically date from after the beginning of industrial fisheries, perhaps most complicating of all challenges in fisheries science may therefore be the fact that for the great majority of stocks, no reference data are available for pre-fisheries population sizes and dynamics. Consequently, the impact of harvesting can only be inferred or estimated, and not fully quantified. Research is heavily dependent on landing data from commercial fishing vessels, and vast discards at sea, illegal fishing activities and unreported landings all greatly affect the reliability

---

<sup>1</sup> Rutter wrote in 1902: “A large fish is worth more on the markets than a small fish; but so is large cattle worth more on the market than small cattle, yet, a stock-raiser would never think of selling his fine cattle and keeping only the runts to breed from. It would be better for the salmon as a species, and therefore better for the salmon industry, if the present minimum net-mesh were made the maximum.” (Rutter 1902, p 134).

and accuracy of fisheries data and complicate its interpretation to aid successful management of stocks (Myers et al. 1997; Watson and Pauly 2001).

The failure to understand the effects of harvesting on population dynamics is reflected in the number of collapsed stocks and declining catches as well as the lack of recovery of stocks after exploitation has stopped (Myers et al. 1997; Hutchings 2000; Petitgas et al. 2010). In some stocks that have been severely diminished by exploitation, recruitment is now so low that it is not sufficient to allow populations to return to their previous abundance levels, even after fishing has been brought to halt (Hutchings 2000). While it is always true that zero spawners will produce zero recruits, more than zero spawners do not always result in viable recruits and for some stocks at least, population dynamics at low levels of abundance seem to be different from those at healthy population sizes. Although conclusive evidence for such compensatory processes is rare (Myers et al. 1995; Liermann and Hilborn 2001), the reasons for an observation of reduced recruitment could further lie in changes in trophic interactions and community structure (de Roos et al. 2006), the age structure of the remnant populations (Berkeley et al. 2004), or increased natural mortality (Swain et al. 2007; Swain 2011).

### **1.2.2 Fishing as a selective force**

Fishing mortality is non-random. The type of gear and mesh size used, depth, season and area targeted affect the catchability of different fish (Swain 1993; Arreguín-Sánchez 1996; Heino and Godo 2002; Kuparinen et al. 2009) and imply that some individuals are more likely to get caught than others. Fisheries are, for example, selective on behaviour, as active fish are more likely to get caught than cautious fish (Heino and Godo 2002; Biro and Post 2008). If such differences exist between sexes, for example differences in mating behaviour or time spent in a high-risk fishing ground, fishing can be selective on sex and result in biased sex ratios (for example in plaice *Pleuronectes platessa*; Solmundsson et al. 2003). Strong schooling fish are most likely to be caught by nets, whilst bold fish are more frequently caught on baited lines (Fernö and Olsen 1994). Depending on depth differently aged fish can be accessed (Swain 1993); fishing on spawning grounds targets solely adult individuals,

whereas catch is irrespective of maturity status on feeding grounds. By varying net mesh size, only the largest or intermediate size classes can be targeted, which can drive selection on growth (Reis and Pawson 1999; Swain et al. 2007). The intricate ways in which fishing acts as a selective force are the subject of decades of study and are discussed in some excellent reviews (Arreguín-Sánchez 1996; Heino and Godo 2002). Here, we focus on a specific aspect of fishing: selection on body size and its consequences for life histories of exploited species.

### **1.2.3 Fisheries-induced change in life history traits**

Selection on body size can affect life history in several ways, which are not mutually exclusive and can be hard to disentangle. Most importantly, the underlying mechanisms of a response can be very different in nature and hard to unravel. Body size is the complex end product of a suite of individual characteristics including life history trade-offs, genetic make-up, and a range of physical factors such as temperature and salinity (Ricker et al. 1978; Conover and Present 1990; Krohn et al. 1997) and biological parameters such as density (Lorenzen and Enberg 2002), food availability (Krohn et al. 1997; Brander 2007) and trophic interactions (Steingrund and Gaard 2005) affecting size at birth, survival, growth and maturation. By definition, increased mortality goes hand in hand with a reduction in population size and associated changes in density-dependent processes and competition for resources. Especially in fish, the most phenotypically plastic of all vertebrates (Rochet 1998), changes in growth and faster maturation may be a phenotypically plastic response to such changes in the environment. In many species, so-called compensatory growth (Policansky 1993) has been observed, where per capita growth increases at low levels of abundance as a plastic response to (harvesting induced) low densities, reduced competition and higher per capita food levels (Policansky 1993; Trippel 1995; Heino and Godo 2002; Engelhard and Heino 2004a). The required size of maturation is then reached at an earlier age.

Amongst others, natural populations of cod (Olsen et al. 2004; Swain et al. 2007), plaice (Rijnsdorp 1993; Grift et al. 2003, 2007), herring (Engelhard and Heino 2004a, 2004b), grayling (Haugen and Vollestad 2001) and salmon (Ricker et al. 1978; Ricker

1981; Hard et al. 2008) have been observed to mature at a younger age and attain smaller body sizes after decades of intense and size-selective harvesting (see for more references Jorgensen et al. 2007).

However, although some of the observed changes can be explained by faster maturation resulting from compensatory growth (Engelhard and Heino 2004a; Hilborn and Minto-Vera 2008), size (Ricker et al. 1978) and growth rate (Swain et al. 2007) actually decreased for several stocks as well, despite favourable conditions. In these stocks, a solely plastic response is insufficient to explain the trends in maturation and body size and genetic adaptation to harvesting offers a potential explanation.

The response of life history traits to such a high, and selective form of mortality like fishing would be expected from life history theory, which states that investments in reproduction are likely to increase at the cost of investments in growth under high predation (Law 1979; Magnhagen 1991; Roff 1991). Harvesting selectively removes older and larger individuals from populations and increases reproductive opportunities for small and early maturing fish. Thereby, fishing enhances fitness for fish that allocate relatively more resources to reproduction at an early age. Provided that growth and maturation traits have a genetic basis, it follows that the frequency of early maturing genotypes will increase under harvesting pressure (Law 2000).

#### **1.2.4 Fisheries induced evolution?**

For evolutionary change induced by fishing to take place, two basic requirements must be met that apply to all occurrences of adaptation in nature: firstly, the trait must be heritable, that is; a proportion of the observed trait variance must result from additive genetic variation (Falconer and Mackay 1996) and second, there must be a mismatch between the fitness of the current trait value and the optimal trait distribution (Hendry et al. 2011). For fisheries-induced selection in particular, this means that selection by harvesting must be of such magnitude that it outpaces possible natural selection acting on the same trait in another direction (Conover 2007; Edeline et al. 2007; Stenseth and Dunlop 2009).

Substantial evidence exists that indicates the presence of heritable genetic variation in life history traits in fishes: the presence of countergradient variation for life history traits over environmental gradients (Conover and Present 1990; Conover and Schultz 1995; Conover and Baumann 2009) and successful selective breeding achievements in the aquaculture industry (Jonasson et al. 1997; De-Santis & Jerry 2007) provide strong evidence for a quantitative genetic basis and heritable genetic variation of life history traits (Law 2000).

A mismatch between population mean values for body size and maturation age, and those optimal under fishing pressure is also evident because of the non-random nature of fishing, targeting – in most cases – larger and older individuals. Similarly, large fish are favoured by natural selection in cases where fecundity increases with age and size (Trippel 1995; Walsh et al. 2006; Fenberg and Roy 2008; Johnson et al. 2011) and when fish can outgrow natural predation risk (Heino and Godo 2002; Conover 2007; Edeline et al. 2007). Additionally, fisheries selection, expressed as a high and selective mortality, embodies a much stronger selective force than natural selection for large size and late maturation (Law 2000). Especially on feeding grounds and in heavily exploited stocks, mortality is of such magnitude that fish are often caught before attaining their reproductive age. In cod the age of 50% maturity was around five years in the mid 1990s (Olsen et al. 2004), but cod were caught as young as three years of age (Myers et al. 1997). In some fisheries, mortality is so high, that it affects all adult size classes, virtually removing the size selective component and it can be hard to distinguish whether selection or “simply” increased mortality is the cause of observed declines (Jackson 2001; Fenberg and Roy 2008; Swain 2011). Even in the absence of a size selective component to fishing, or without immature individuals being caught, a decrease in maturation age would still be expected from life history theory, since increased mortality translates to a reduction in lifetime reproductive success and favours an early onset of the reproductive lifespan (Roff 1992).

Less clear than determining a mismatch between optimal and observed trait values, is to establish whether observed changes match selection differentials imposed by

fishing (Kuparinen and Merilä 2007; Law 2000, 2007). The nature and timing of fisheries harvesting greatly affect the intensity and direction of selection it embodies, and plasticity and compensatory growth can obscure selection for reduced growth and small size (Rochet 1998; Heino and Godo 2002; Nusslé et al. 2011). In relation to this, growth at early ages in fish is to a large extent dependent on environmental variables and can determine size later in life, thereby either obscuring or reinforcing evolutionary trends (Letcher et al. 2011). As many species have indeterminate growth (Heino and Kaitala 1999), reductions in size can also be the result of a truncated age structure (Hsieh et al. 2006) and do not necessarily represent phenotypic shifts in size-at-age. Correlations between life history traits further complicate the inference of selection differentials as it is hard to distinguish between traits under selection and co-evolving traits (Heino and Godo 2002; Funk et al. 2005; Munch et al. 2005). Harvesting is seldom the only selective force working on exploited populations over time and other (environmental) factors need to be considered that could explain observed trends in life history traits and phenotypes of exploited stocks (Daufresne et al. 2009).

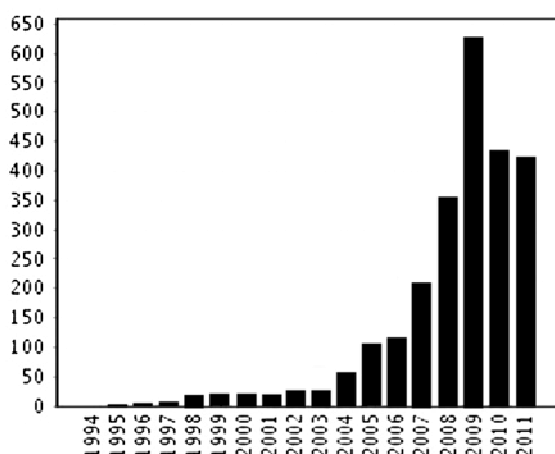
Such complexities make it difficult to reliably estimate an evolutionary response to fishing and to compare quantitative estimates of selection differentials from fishing with observed phenotypic shifts (Law 2000, 2007). Fortunately, an increasing body of research is dedicated to this matter and realistic estimates of heritability and quantitative genetic parameters in wild populations are becoming more readily available (Jonasson et al. 1997; Haugen and Vollestad 2001; Funk et al. 2005; DiBattista et al. 2009; Nusslé et al. 2009) although direct links between estimates of quantitative parameters and fishing pressure are still scarce.

### **1.2.5 Discriminating between evolutionary change and phenotypic plasticity**

Evidence of fisheries induced changes of life history traits dates from as early as the 1950s (Miller, 1957) but has been largely ignored by the scientific community until empirical studies demonstrated declines in growth, age and size of maturation in harvested stocks a few decades later (Handford et al. 1977; Bigler et al. 1996; Ricker 1981; Rijnsdorp 1993). Even then, a proper scientific debate did not ensue (but see

Stokes et al. 1993; Rochet 1998) until the 21<sup>st</sup> century (Browman et al. 2000; Law 2000; Kuparinen and Merilä 2007; Zhou et al. 2010). The field of “Darwinian fisheries science” as coined by Conover (Conover 2000) has become a rapidly growing field of study: a European research network dedicated entirely to Fisheries-Induced Evolution (FinE: [www.iiasa.ac.at/research/EEP/fine/home.html](http://www.iiasa.ac.at/research/EEP/fine/home.html)) began in late 2007 and a brief search on Web of Knowledge for “Fisheries (-) induced evolution” and “Darwinian fisheries” yields 112 of 142 hits published from 2007 onwards, compared to no hits before the 1990s (Fig. 1.1).

This increasing interest in selective effects of fisheries, and in particular its possible evolutionary consequences, is explained by an increasing concern over the potential negative effects of evolutionary change on harvested stocks and yield (Browman et al. 2000; Law 2007; Fenberg and Roy 2008). If selective fishing has evolutionary consequences on the reproductive and physiological characteristics of harvested species, understanding these effects is of critical importance, since evolutionary change in life history traits may irreversibly affect a fisheries yield. In the long run (Law 2000, 2007) genetic erosion may lead to a loss in adaptive potential (Smith et al. 1991; Hauser et al. 2002; Hoarau et al. 2005); changes in life history traits may be hard to reverse if strong selection for large body size is lacking.



**Fig. 1.1.** Citation report of the number of studies citing 142 studies on: “fisheries(-) induced evolution” or “Darwinian fisheries”. Created from Web of Knowledge: <http://apps.webofknowledge.com>.

Further debate focuses on the strength of the – mostly phenotypic – data in providing evidence for fisheries induced genetic change (Browman et al. 2000; Hilborn 2006; Hilborn and Minte-Vera 2008; Kuparinen et al. 2008).

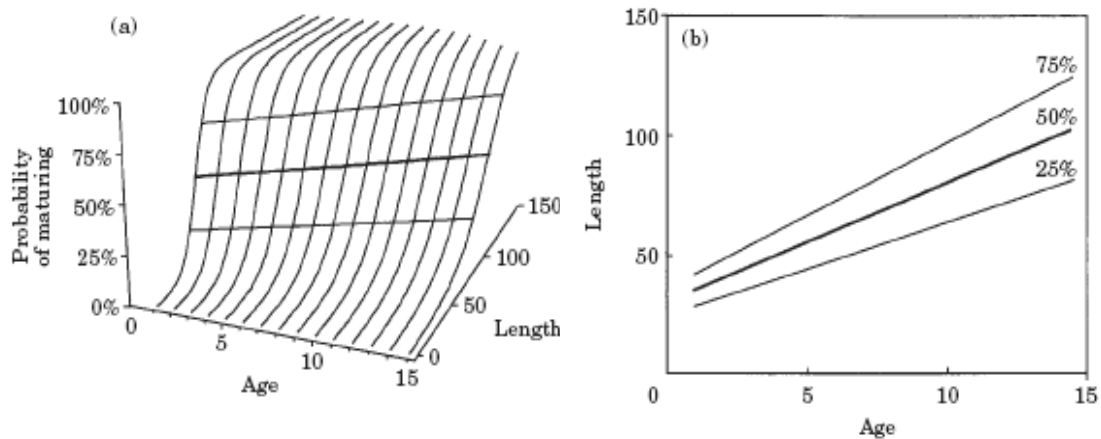
One of the earliest and still most convincing examples on fisheries induced evolution comes from Ricker



et al. (Ricker et al. 1978) who demonstrated reductions in size of Pacific Salmon *Oncorhynchus gorbuscha* to be associated with fishing pressure. As these fish almost always mature at 2 years of age, plastic components affecting maturity could be eliminated as a cause of reduced growth, and temperature and salinity trends could explain only part of the observed declines in size. Furthermore, feeding conditions had actually improved over the study period, and a plastic response to environmental changes would have been expected to cause directional change in size in the opposite direction from what was observed. These factors altogether strongly suggest there was a genetic response to harvesting, causing reduced growth. Ricker (1981) demonstrated comparable trends in other species of salmon and considerable work has been done on the effects of fishing on salmon life history traits since then, demonstrating heritable components to body size (Jonasson et al. 1997; Smoker et al. 2000; Funk et al. 2005) and confirming substantial fisheries selection (reviewed in Hard et al. 2008), although, like elsewhere, only phenotypic data is available to infer evolutionary change.

#### **1.2.6 The use of probabilistic reaction norms for inferring genetic change**

Examples such as Ricker et al. 1978 are rare and by far the majority of evidence so far is provided by studies using probabilistic maturation reaction norms (PMRN, (Stearns and Koella 1986; Heino et al. 2002), which have been very useful in clarifying the most important drivers behind observed life history shifts (Dieckmann and Heino 2007). The method plots the probability of maturation as a function of age and size (Fig. 1.2). By focusing on the maturation process itself, the approach attempts to disentangle the environmental processes affecting growth and survival from the genetic component determining whether attaining a certain age or size makes maturation likely (Rijnsdorp 1993; Heino et al. 2002). The method has the advantage of relatively straightforwardly visualising trends in maturation independent of environmentally-induced variation in growth, using data readily available from survey and catch data of most exploited stocks. It has been criticised however, for being too reliant on maturation being a function of just age and size and requiring trait variation to be due to environmental factors only (Kuparinen and Merilä 2007; Morita et al. 2009).



**Fig. 1.2.** (a) Reaction norms for age and size at maturation describing the probability of immature individuals to mature as a function of size and age. (b) Projection of the 50% midpoints and width of the range of reaction norms on a plain, results in a 2-dimensional representation of the probability of maturation at a certain age-size combination. Figure from: Heino et al. 2002.

If this requirement is not met, PMRNs may be more representative of genetic variation in other traits affecting maturation than of phenotypic plasticity in maturation age and length. Exactly this shortcoming of the reaction norm approach has received criticism from peers: it assumes that maturation is a mere function of current age and size, and that genetic components of traits such as body condition and growth rate leave the maturation probability unaffected (Kuparinen and Merilä 2007). However, in salmon for example, there is evidence that recent growth history is a greater indicator of maturation probability than age or size (Morita and Fukawaka 2006). Barot et al. (2005) found similar reaction norms for different American plaice stocks with different fisheries selection pressures, indicating factors other than fisheries' selection contributed to the norm in some way. Transplant experiments with charr (*Salvelinus leucomaenis*) also demonstrated the PRMN has a plastic component (Morita et al. 2009). In Atlantic cod, *Gadus morhua*, some of the change in age and size of maturation suggested to be a response to harvesting using a reaction norm approach (Olsen et al. 2004) took place in cohorts that were present before a change in exploitation regime occurred. Such an observation implies that environmental change is a more likely explanation for observed changes than genetic adaptations to harvesting (Kuparinen and Merilä 2007; Law 2007).

Multidimensional reaction norms offer a promising means to surmount some of these problems and three dimensional reaction norms have been obtained amongst others for cod (Baulier et al. 2006) sole (Mollet et al. 2006) and plaice (Grift et al. 2007; Kraak 2007), and included body condition and/or temperature as variables. These studies found that the third dimension could explain part of the trends initially attributed to evolutionary effects of fishing but that this contribution was small (e.g. 3-8% in Grift et al. 2007), and a large part remained unexplained. In theory, the reaction norm model can be extended with an unlimited number of axes, including any environmental or physiological parameters on which sufficient data are available. Such multidimensional models are increasingly hard to interpret, however, and reaction norms with more than three dimensions are, to our knowledge, currently nonexistent.

### **1.2.7 The missing link: conclusive evidence of genetic change**

It is not that critics consider the process of fisheries-induced evolution (FIE) to be unlikely; they question the claim that phenotypic data provide conclusive evidence on evolutionary change by selective harvesting (Kuparinen and Merilä 2007; Law 2007; Morita et al. 2009) and the magnitude of FIE as claimed by some (Conover and Munch 2002; Jorgensen et al. 2007). There is ample evidence demonstrating life history changes in exploited species (e.g. reviewed in Jorgensen et al. 2007), but at best very little quantitative or direct molecular genetic evidence is available to support hypotheses on the impact of selective fishing on the genetic composition of populations (Kuparinen and Merilä 2007; Law 2007).

Studies demonstrating a genetic response to fishing are rare for obvious reasons: disentangling genetic effects from phenotypic plasticity can be notoriously difficult, as described in 1.2.5. Secondly, numerous selection pressures operate simultaneously on wild populations, hindering the unequivocal association of a selection pressure with specific phenotypic shifts or individual genetic loci. Thirdly, life history traits, and body size in particular, are quantitative traits of which the genetic basis is still poorly understood. Numerous loci are likely to affect trait distributions due to their polygenic nature. Consequentially, a substantial amount of

non-additive genetic variance contributes to variance of polygenic traits, and phenotypic change is typically much larger than measurable change at individual loci (Falconer and Mackay 1996).

One notable exception is a recent study by Jakobsdóttir et al. (Jakobsdóttir et al. 2011) who examined allele distributions at the *Pantophysin* locus (*Pan 1*) in a temporal data series (1948 – 2002) of Icelandic cod. Their analyses showed a strong decrease in the frequency of the *Pan 1*<sup>BB</sup> genotype over time, which also had significantly higher frequencies in older fish within single cohorts. Their data suggest strongly that this allele was associated with low fishing pressure and late maturing fish, and demonstrate the value of complementary research strategies; using temporal series of both phenotypic and genetic data at neutral and putative candidate loci. Although the authors do not claim it as such, this study is to the best of my knowledge the only available study to demonstrate fisheries induced changes at a specific locus.

There is an urgent need for studies providing conclusive quantitative- (i.e. matches of observed phenotypic changes with likely heritabilities and fishing selection differentials) and qualitative genetic evidence (i.e. molecular signatures of selection) that can examine whether fisheries selection on body size is responsible for observed trends, and whether this is the result of evolutionary change or phenotypic plasticity. The present study responds to this call by quantifying the selective response to harvesting in experimental populations of Trinidadian guppies (*Poecilia reticulata* Peters). By means of a size selective breeding regime of known intensity and over multiple generations, closely monitoring both phenotypic shifts in life history traits and genetic dynamics, we test directly the hypothesis that harvesting causes genetic change, and whether it is possible to identify potential candidate loci.

## **1.3 SELECTION EXPERIMENTS AND THE STUDY OF ADAPTATION**

### **1.3.1 Advantages and limitations to the use of selection experiments**

Selection experiments allow genetic and environmental aspects of adaptation to be experimentally investigated (Fuller et al. 2005; Conover and Baumann 2009). Experiments form one of the only means to directly test evolutionary hypotheses; to couple a selective agent directly to its response, removing concealing effects of environmental variation and other selective pressures that are likely to be present in the wild. They allow the magnitude of the response to selection in traits of interest and correlated traits to be quantified. Hereby, the extent to which genetic correlations shape and constrain the evolution of traits can be revealed. Furthermore, experiments are the only means to confirm an actual heritable component to any trait by eliminating potentially confounding environmental variables and trends (Conover and Baumann 2009).

There is a large body of scepticism towards experimental approaches to the study of adaptation in the wild. Simplification of natural systems in laboratory environments could affect heritability estimates and alter the nature of the response to selection compared to natural environments (Holloway et al. 1990; Hoffmann and Merilä 1999 but see for example the studies by Weigensberg and Roff 1996; St. Juliana and Janzen 2007). Any gene x environment interactions present in the wild but not in the laboratory will further distort  $h^2$  estimates and complicate the translation of experimental results to the wild (Hoffmann and Merilä 1999; Ellegren and Sheldon 2008). Furthermore, there is often a taxonomic bias towards short-lived species with fast reproductive cycles, which cannot necessarily be compared directly with species of interest in the wild. For example, such fast-reproducing species are often characterised by large census and effective population sizes, which according to Fisher's Fundamental Theorem (Fisher 1930), should elevate the adaptive evolutionary potential compared to species with smaller population sizes.

Despite these limitations, selection experiments have been immensely fruitful in providing insights into the nature of- and correlations between traits responding to

selection (Swallow et al. 1998; Garland et al. 2002) as well as in some major evolutionary processes, such as the strength of genetic correlations (Beldade et al. 2002); the generality of trade-offs (Rose et al. 2005); the relative importance of phenotypic plasticity and genetic composition (Scheiner 2002) and the identification of genomic regions under selection through quantitative trait locus (QTL)-mapping (Rand et al. 2010). However, examples of artificial selection experiments combining quantitative and molecular approaches in non-model organisms are still rare (Fidler et al. 2007).

### **1.3.2 Experimental studies on fisheries-induced evolution**

At least four previous experimental studies have investigated fisheries-induced evolution, though none of these studies utilised molecular genetic markers to explore the nature and extent of genetic change.

Silliman (1975) used an admixture population of *Tilapia mossambicus* (*Oreochromis mossambicus*) and harvested 10-20% of individuals every few months, using body thickness as selection criterion, by harvesting only those fish that were capable of swimming through a defined width between two rods. It took around six generations to observe a strong reduction in yield and male growth.

Edley and Law (1988) performed repeated episodes of harvesting on different size classes of mixed clonal lines of *Daphnia magna*. Every few days, 40% of individuals in the vulnerable size classes were culled and an increase in frequency of slow growing clones in large-harvested lineages, compared to small harvested replicate populations, was recorded. Additionally, divergence in life history traits other than growth rate (age and size at maturation) was observed.

Thirdly, an artificial selection experiment on Atlantic silversides *Menidia menidia* (Conover and Munch 2002) demonstrated a very rapid response of life-history traits to four generations of selection on adult body size. Not only body size responded to selection but a range of life history, physiological and behavioural traits co-evolved and changed significantly as a result of the selection regime (Walsh et al. 2006).

Finally, natural selection experiments on Trinidadian guppies (Reznick et al. 1990, 1997) showed rapid and strong evolutionary responses in life history parameters to relaxed predation pressure (further explained in 1.4.3). The guppy experiment is of particular interest here as it is one of the strongest examples known to date of life history evolution due to size-selective harvesting in the wild. There is a large body of previous- and ongoing work on adaption of guppies to environments with different size-selective predation regimes (Magurran 2005), making these fish an informative model for studying FIE and the ecology and evolution of life-history strategies in general.

## **1.4 THE TRINIDADIAN GUPPY, *POECILIA RETICULATA*, AS A MODEL SPECIES**

### **1.4.1 *Poecilia reticulata* as a research model**

The Trinidadian guppy (*Poecilia reticulata* Peters, Poeciliidae, Cyprinodontiformes, Actinopterygii) is a small, lecithotrophic, viviparous freshwater fish native to Venezuela, Guyana, Suriname and the Caribbean islands of Trinidad and possibly Tobago. On Trinidad, it is by far the most widespread and abundant fish species, and it can be found in most freshwater bodies on the island (Haskins and Haskins 1951; Magurran and Phillip 2001). It has been introduced throughout the world for mosquito control (Courtenay and Meffe 1989) and can now be found throughout the Caribbean and the Americas, parts of Africa, Southeast Asia and Oceania. Escapees from the aquarium trade have even succeeded in establishing successful populations in warm effluent waters of power plants in temperate countries (Fig. 1.3).



**Fig. 1.3.** Worldwide distribution of *Poecilia reticulata* showing native range (black) and invasive distribution (dark grey). Native distributions are taken from Magurran (2005), invasive distributions from: [www.fishbase.org](http://www.fishbase.org) and may be incomplete. With the exception of Russia (Moskou only) distribution of *P. reticulata* may vary within countries, which is not indicated on the map. Inset shows the island of Trinidad with the three Northern river drainages and the Tacarigua River (Ta).



Although the species has gained tremendous popularity through domestication and breeding of strains for the aquarium trade, for which a variety of guppy strains is being used worldwide, it is the original Trinidadian guppy that still is by far the most popular subject of research. This is a result of the fact that much of the geographic variation in behaviour, morphology, and life history traits can be related to a key selective force (predation) and can be interpreted in relation to corresponding fitness variation (see 1.4.3).

The earliest known papers on the species date back to the second half of the 19<sup>th</sup> century (Peters 1859; Günther 1866) but its popularity for evolutionary studies began over 50 years later with work on sperm competition (Schmidt 1919; Winge 1922a; Winge 1937) and sexual selection (Haskins & Haskins 1949). However, it was not until the work of Endler and Reznick, who disclosed the importance of predation intensity on morphology and life history (Endler 1980, 1988; Reznick and Endler 1982), that the guppies' popularity as a model organism truly exploded. Since then, the species has gained increasing popularity with around 95% of papers being published since 1990. The species is a popular model organism for the study of life history (Reznick et al. 1992; Reznick et al. 2001a), sperm competition and sexual selection (Matthews et al. 1997; Evans and Magurran 2001), population genetics (Shaw et al. 1994; Carvalho et al. 1996); conservation biology (van Oosterhout et al. 2007a), senescence (Reznick 1997; Reznick et al. 2001b; Reznick et al. 2004), behaviour (Magurran et al. 1995; Croft et al. 2009), parasitology (Cable and van Oosterhout 2007) and fisheries-induced evolution (Reznick and Ghalambor 2005a), to name but a few out of the breadth of fields that have benefitted from the extraordinary research qualities of this small fish.

The reasons for such attention are multiple. A good model species can be studied both in its natural habitat and in manipulated environments or laboratories. It responds quickly to manipulations in a measurable way, is easy to breed and has sufficiently short generation times to allow studies over multiple generations within the timeframe of standard research projects. Ideally, a model species' biology is well documented in the literature and it can be found across diverse ecological conditions

in the wild within a small geographical distance, allowing for the assessment of different environmental factors on the species' characteristics. Guppies have all such attributes (Magurran 2005) and in addition, have several of the specific qualities that are required for specialist models of specific fields of research: their sexual polymorphism in colour and heritable colour pattern variations make them ideal for studying sexual selection (Houde 1997); the possibility of controlled mating and artificial insemination (Evans et al., 2003), combined with the ability of females to store sperm (Winge, 1922a, 1937) and multiple paternity within broods provide a convenient setup for testing hypothesis on sperm competition and pre- and post copulatory female choice (Evans and Magurran, 2000); it is possible to tag and mark fish individually for monitoring of social interactions and survival in the wild (Reznick et al. 1996; Croft et al. 2003a, 2003b); their tolerance to a range of temperatures, water qualities and toxic substances, makes guppies a popular subject in toxicological studies (Haubruge et al., 2000; Magurran and Phillip, 2001). Furthermore, guppies show a range of interesting behaviours as reproductive, feeding and predator avoidance strategies and are extremely popular in behavioural research (Magurran and Seghers 1991; Godin and Briggs 1996; Laland and Williams 1997). Last, but by no means least, guppies owe their popularity to the fact that throughout the rivers in the Northern Mountain ranges of Trinidad, marked differences in phenotypic variation at morphological, behavioural and life history traits is repeatedly associated with variation in predation levels between stream parts.

#### **1.4.2 Population distribution; phylogenetic history and connectivity**

The Trinidadian mountains host many parallel rivers running down into two main drainages: the Caroni drainage to the West of the island and the Oropuche to the East (Fig. 1.3). These drainages have been separated for at least 600,000 years (Carvalho et al. 1991) and are host to highly divergent clades of guppies (Carvalho et al. 1991; Fajen and Breden 1992; Breden et al. 1999; Alexander et al. 2006; Schories et al. 2009; Willing et al. 2010). In addition, seven notable rivers run down the Northern slopes of the mountains directly into the sea. These were initially believed to be colonised from the Caroni drainage (Carvalho et al. 1991; Fajen and Breden

1992), but more recent analyses suggest that these comprise a third phylogenetic lineage (Suk and Neff 2009; Willing et al. 2010). Guppies can be found mostly in smaller pools, streams and tributaries, rather than in fast flowing or deep bodies of water throughout the three drainages with exception of some of the Northern rivers, which guppies have failed to colonise, most likely due to their isolation rather than unsuitability of habitat (Magurran and Phillip 2001).

Genetic analysis has revealed a clear pattern of differentiation between populations within streams, between rivers and between drainages (Carvalho et al. 1991; Shaw et al. 1991; Shaw et al. 1994; Barson et al. 2009; Willing et al. 2010). Within rivers, upland populations typically depict low levels of local genetic diversity, whereas fish further downstream have more genetic variability and are often an admixture of populations upstream as well as neighbouring rivers. It follows from this that the highest levels of differentiation between rivers can be expected between upland populations; as supported by genetic data (Shaw et al. 1991; Crispo et al. 2006; Barson et al. 2009). Limited gene flow upstream due to barrier waterfalls and strong currents flushing fish downstream during heavy rainfall, explain this source-sink pattern of differentiation (Shaw et al. 1991; Crispo et al. 2006; Barson et al. 2009). The strongest divergence is observed however between drainages (Carvalho et al. 1991; Willing et al. 2010).

The parallel population distributions of guppies in multiple rivers provides a unique system for studying effects of stochastic differentiation, colonisation history and ecological effects on dispersal, population distribution and persistence (Magurran 2005 and references therein). Furthermore, and of particular relevance to the present study, is the fact that predation levels vary between up- and lowland river parts and have shaped the life history of guppies repeatedly, and in similar ways, throughout the river basins.

#### **1.4.3 Predation and life history strategies**

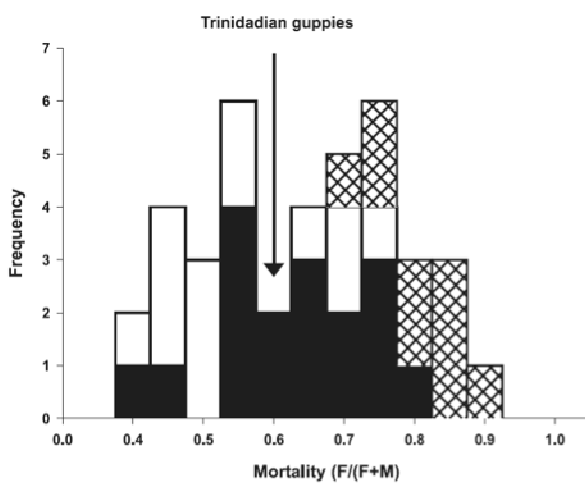
Throughout the lower parts of the rivers, guppies share their habitat with different predatory fish communities. Predators such as *Crenicichla alta*, *Hoplias malabaricus*,

*Aequidens pulcher* and *Astyanax bimaculatus*, can be found in high numbers and are fervent guppy predators (Magurran and Phillip 2001), in particular of adult individuals. Further upstream, the distribution of these predatory fish is limited by waterfalls, which form barriers to their dispersal (Endler 1978). Predation on guppies in upper parts of streams is mainly by the killifish *Rivulus hartii* but only large individuals of this species are capable of feeding on adult guppies and *Rivulus* is mainly a threat to smaller and juvenile guppies only (Mattingly and Butler 1994). Initially, it was thought that the predation difference between the different communities results from a difference in predation intensity as a function of body size; mortality is notably higher for larger individuals in “high predation” sites, but not for juveniles. Later analyses showed that this is not the case, but an overall difference in mortality characterises the two predation regimes (Mattingly and Butler 1994; Reznick et al. 1996).

Although the importance of predation on life history evolution is extremely well studied, surprisingly little is known about the feeding patterns of the different predatory fish and which species actually prey on *P. reticulata* as a major part of their diet (Haskins et al. 1961; Endler 1987; Fraser et al. 1995). The impact of avian predators and invertebrates is also poorly quantified (Endler 1983; Rodd and Reznick 1991). What is evident however, is that predation intensity is much lower, both on adults and juveniles, in the upland communities compared to populations further downstream (Mattingly and Butler 1994; Reznick et al. 1996). This is also true for populations in the rivers North of the mountains, though the species composition of the predatory community is different; the main piscivores are gobies (*Eleotris pisionis*) (Reznick and Bryga 1996).

Endler (1980, 1983) first investigated the role of predation in the evolution of guppies in contrasting habitats, demonstrating colour variation to be the result of sexual selection for brightly coloured males and natural selection for colours that do not stand out from their surroundings in heavily predated areas. Subsequent experiments demonstrated that predation level does not only affect male colour patterns, but affects a whole range of life-history traits in a manner in agreement

with life-history theory (Reznick and Endler 1982; Reznick et al. 1996). Where mortality is high, individuals produce more offspring more frequently and start reproduction at an earlier age, whereas under low predation pressure – where individuals can be expected to have a longer reproductive lifespan – relatively more resources are allocated to growth as opposed to reproduction. The differences persisted over two generations of breeding under common garden conditions, confirming these consistent differences to have a genetic basis (Reznick 1982; Reznick and Bryga 1996). No less than 50 traits, ranging from life history traits such as reproductive allotment and timing of maturation, to sex ratios, colour patterns, behaviour strategies and levels of genetic variation, have been repeatedly found to co-vary with predation regime in different rivers (Endler 1995; Reznick et al. 1996; Magurran 2005) as well as growth rate (Arendt and Reznick 2005), parasite resistance (Cable and van Oosterhout 2007) and morphometrics (Palkovacs et al. 2011).



**Fig. 1.4.** Frequency distribution of mortality values of guppies and exploited marine species, expressing fishing mortality  $F$  as a proportion of total mortality  $F+M$ . Fish species are divided over three categories: pelagics (open bars), heavily exploited groundfish (cross-hatched) and other species (black). Data for guppies comes from Reznick et al. 1996, marine species from Mertz and Meyers 1998. Figure from: Reznick and Ghalambor 2005a.

Using mark-recapture data from previous work (Reznick et al. 1996), Reznick and Ghalambor (Reznick & Ghalambor 2005a) expressed differential mortality of high- and low predation populations as a proportion  $F$  of total mortality ( $F+M$ ), a common expression of fishing mortality in fisheries science, and showed that this mortality falls in the middle of the range of fishing mortality levels experienced by marine exploited fish species (Fig. 1.4). Direct comparisons of harvesting-induced life history shifts of marine species and heavily

predated guppy populations are therefore possible and seem relevant.

#### **1.4.4 Transplant experiments**

Perhaps most convincing in demonstrating the role of predation in life history evolution, are a series of natural selection experiments where “high predation” guppies (originating from an environment with prominent guppy predators such as *C. alta*) were introduced to a “low predation” environment (with only the weak predator *R. hartii* present) in a tributary of the Aripo River (Endler 1980) and the Guanopo River (El Cedro) (Reznick and Bryga 1987). Over the course of 11 years and 7.5 years respectively, life histories of the introduced fish changed rapidly and significantly, increasingly resembling low predation guppies over time: having large body size with late onset of maturation; producing fewer, large offspring; males displaying bright colour patterns. These differences persisted under several generations of common garden rearing, confirming a genetic basis, rather than phenotypic plasticity driving the observed shifts (Reznick and Bryga 1987; Reznick et al. 1990). In a third, more recent introduction experiment (1996), guppies from a high predation site in the Yarra River on the Northern Slopes, were introduced into a low predation environment of the neighbouring Damier River (Karim et al. 2007). The introduced fish colonised high predation sites further downstream and over less than ten years, developed colour patterns and life history traits different from their ancestral population, as well as different between habitats with different predation regimes in the Damier river (Karim et al. 2007; Gordon et al. 2009).

#### **1.4.5 Environmental differences between upland and lowland habitats**

Predation pressure has repeatedly demonstrated to be of great importance in shaping the life history evolution of guppies (see 1.4.3-1.4.4) and makes guppies a useful model to study effects of harvesting predation on exploited fish species (Reznick & Ghalambor 2005a). However, predation is not the only factor that differs between up- and downstream river parts and like for marine fish, environmental factors other than predation may affect growth in guppies.

Although physical properties of the water (temperature, water quality, pH) are no different between sites, canopy cover and population density differ can vary considerably within and between rivers. In areas with low canopy cover, light

availability is significantly higher than in areas with dense forest cover result in elevated primary productivity of algae – the primary food source for wild guppies. Consequently, food availability differs between sites, which in turn affects guppy growth (Grether et al. 2001; Reznick et al. 2001a; Schwartz and Hendry 2010). For example, algal availability and canopy cover can explain the majority of variation in growth rates between different low predation sites (Grether et al. 2001). As canopy cover typically is lower at high predation localities, the early maturation and small body size of guppies at such localities may thus as well be partially an effect of compensatory growth resulting from higher food availability.

Guppy population density is another factor that varies greatly between sites, although in a much more erratic pattern than vegetation. In terms of biomass, low predation sites seem to have higher densities of guppies (Reznick et al. 2001a) but greater schooling at high predation sites may affect density, resource availability and competition locally (Rodd and Reznick 1997). These complex dynamics are still poorly understood (Magurran 2005) but in general, resource availability seems to be higher in open streams and high predation sites leading to faster growth rates and larger asymptotic body size (Reznick et al. 2001a). Additionally, parasite load and resistance to parasites (van Oosterhout et al. 2003, 2007b) is higher in high predation environments and may affect growth and survival. Furthermore, year to year variation is considerable due to seasonal variation (Reznick 1989), flooding (Grether et al. 2001) and external influences such as human disturbance, which can affect population dynamics in a dramatic though unpredictable fashion (Magurran 2005; Schwartz and Hendry 2010).

#### **1.4.6 Predator introduction experiment**

To circumvent such co-varying environmental factors that could explain part of the changes observed after the transplant experiment, *C. alta* was introduced into a river stretch of the Aripo River that was previously only inhabited by guppies and *R. hartii* (Reznick 1997). All other environmental factors thus remained equal except predation regime, which increased after the introduction of the predatory fish. After five years, fish in the introduction size were intermediate between control sites up-

and downstream for both age and size of maturation (males) and first parturition (females), although only female age at first parturition was significantly different from the low predation control and the authors fail to report whether these data are obtained from wild-caught fish or from their offspring, reared under common garden conditions (Reznick 1997; Reznick and Ghalambor 2005a). This lack of a response may be evidence of the importance of environmental factors other than predation in determining guppy life history. The experiment is ongoing however and, when compared to the earlier transplant experiments (Reznick and Bryga 1987; Reznick et al. 1990), seems indicative of an ongoing evolutionary response to increased predation. In light of our own experiments, this work is of particular interest as it directly quantifies the effect of increased predation on life history traits. Quantification of such an effect in the wild in this way is, to our knowledge, currently non-existent.

#### **1.4.7 *Poecilia reticulata* as a genetic and genomic model**

The body of research dedicated to unravelling the eco-evolutionary dynamics of Trinidadian guppies and the contributions this work has made to our understanding of life history evolution and natural selection in general is impressive. However, the genetic basis of the traits studied is thus far poorly understood and *P. reticulata* as a genetic and genomic model seems to be still in its infancy, although a significant amount of genetic data has been published in recent years (e.g. Dreyer et al. 2007; Willing et al. 2010; Fraser et al. 2011).

Early work using genetic markers focused on clarifying population structure using allozymes (Carvalho et al. 1991; Shaw et al. 1994); phylogeography of guppies in Trinidad has been further studied more recently using microsatellites (Suk and Neff 2009) and single nucleotide polymorphisms (SNPs) (Willing et al. 2010) as is discussed in 1.4.2. Microsatellite development (Becher et al. 2002; Watanabe et al. 2003; Olendorf et al. 2004; Paterson et al. 2005; Shen et al. 2007a) has greatly aided the study of population genetics (Barson et al. 2009; Suk and Neff 2009) sexual selection and inbreeding (van Oosterhout et al. 2003) and allowed for paternity analyses to be used in mate choice studies (Evans and Magurran 2000; Becher and



Magurran 2004). These and other markers also enabled the development of modest linkage maps (Khoo et al. 2003; Watanabe et al. 2004; Shen et al. 2007b; Tripathi et al. 2009a) giving insights into chromosome number (23) and marker distributions, as well as information on physical correlation between traits and the location of the sex locus and sex-linked traits (Tripathi et al. 2009a, 2009b).

However, knowledge on functional genetic variation has, with few exceptions (Brooks 2000) had to wait for the more recent advances in genomic research that allows investigation of the genetic basis of the traits that have been studied for so long. An extensive database of 18,000 expressed sequence tags (ESTs) was developed in 2007 (Dreyer et al. 2007), marking the onset of the genomic age in guppy research. SNP polymorphisms from 224 of these EST-linked markers and a further 819 BAC-linked markers were subsequently used to map colour pattern (size, area, intensity, hue and centre of mass position) and body characteristics (measurements of- and ratios between width and length of morphometric reference distances) of male guppies (Tripathi et al. 2009a). These authors located putative genomic regions for selection in the wild and made a major step forward towards identifying genes that determine male colouration and body size. Already this work has generated valuable insights into functional genetic variation in the wild (Willing et al. 2010), the genetic basis of important traits such as body size (the present study), orange preference by females (Watson et al. 2011) and spine deformities (Gorman et al. 2011). Recently, a full transcriptome of male and female guppies has been sequenced using 454 GS FLX technology (Fraser et al. 2011); candidate genes and polymorphisms for selection were identified, as well as insights gained into (sex specific) gene expression variation (Fraser et al. 2011). Furthermore, the sequencing of the entire guppy genome is expected to commence in 2011 (Dreyer, pers. comm.).

With such major molecular advances, the Trinidadian guppy joins only a handful of species for which we have such well documented knowledge on life history, behaviour and eco-evolutionary- and population dynamics in the wild, as well as the amenability to experimental manipulations and access to extensive genomic resources.

## **1.5 AIMS OF THE STUDY AND OUTLINE OF EXPERIMENTS**

### **1.5.1 Aims and objectives**

In this thesis, artificial selection experiments on the Trinidadian guppy are used to investigate the genetic basis of a response to size selective harvesting, to contribute to the understanding of life history evolution in guppies, fisheries-induced evolution and the genetic basis of body size variation in fish. Specifically, this thesis is aimed at answering the following question:

*Can directional size-selective harvesting cause detectable genetic change and phenotypic shifts in life history traits of captive populations of Trinidadian guppies over few generations?*

This work addresses the above question by means of four explicit objectives:

1. To quantify the correlated responses of life history traits to size-selective harvesting in guppies
2. To characterise the rate at, and direction in which these traits may change under strong directional selection
3. To monitor the genetic dynamics of neutral microsatellite loci in small, captive breeding fish populations
4. To identify polymorphisms in putative candidate genes associated with phenotypic shifts and a selection response

### **1.5.2 Underlying rationale**

As explained earlier, substantial advances have been made in recent years in our understanding of the circumstances that promote rapid evolution of life history traits in fishes, the rate at which human harvesting can cause phenotypic shifts in commercially exploited species and the genetic basis of adaptation in the wild. However, a conceptual framework in which the knowledge from these advances can be combined to determine the genetic basis of rapid phenotypic shifts in response to human exploitation and disturbance is still lacking. The challenges herein for wild

populations lie in the reliable quantification of selection differentials, which often vary over time, and secondly in convincingly establishing a causal relationship between a selective agent and observed trends (Law 2000, 2007). Supporting such a relationship with an associated genetic trend, in particular at specific loci of known function, is even more challenging and we know of only one example of a study that has succeeded in doing so in marine fish (Jakobsdóttir et al. 2011).

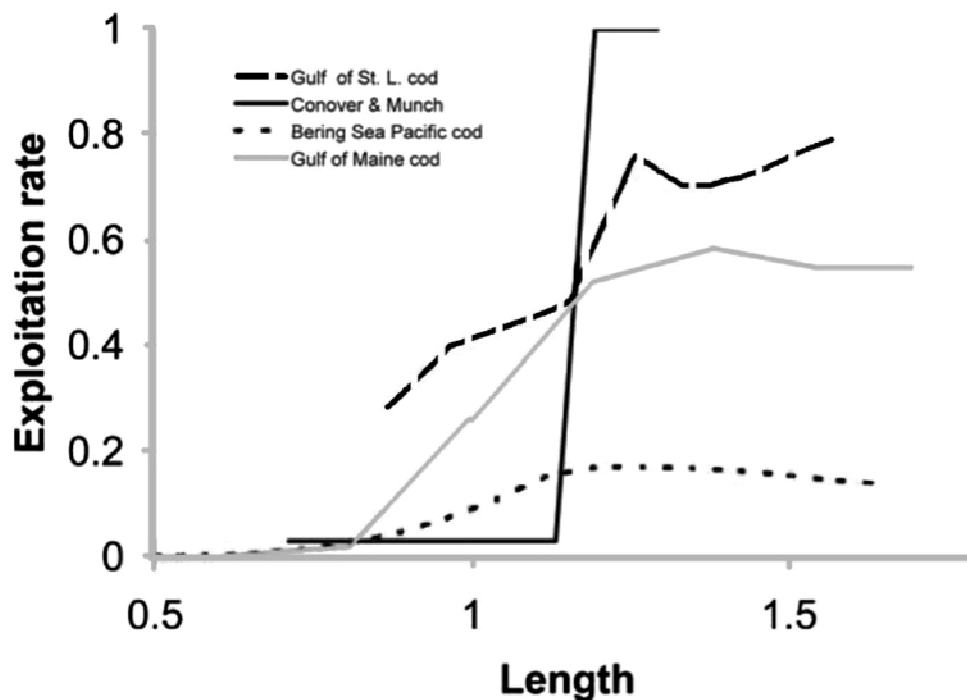
Artificial selection experiments, as used in the present study, maximise control and minimise environmental variation (Fuller et al. 2005). Thereby, they provide a strong method to simplify the complexities involved in disentangling different selective pressures, environmental and genetic change in the wild. Using two different selection regimes of known intensity and direction, in conjunction with neutral microsatellite markers and candidate genes for which substantial support suggests a role in body size (e.g. Tripathi et al. 2009a), the quantitative- and molecular genetic components of a response to harvesting can thus be quantified. Replication within each selection regime will further provide insights into the generality of the response observed.

We have chosen the Trinidadian guppy as a model for its ease of breeding and experimental manipulations and because it is one of a very few species for which the effects of ecology and natural selection on life history have been studied in detail and for which a suit of genetic and genomic tools is available (see 1.4). It is truly a “non model-model organism” and one of the best-known examples of life history evolution in the wild. However, information on the genetic basis of the distinctive traits characterising fish from high and low predation environments is lacking.

This work represents the first time that selection on body size in guppies, as so well-documented to be a major driver of differentiation of wild guppy populations, has been mimicked under controlled conditions. Hereby, the present study contributes directly to our understanding of the genetic basis of life history traits and genetic dynamics of small captive populations. Comparing the outcomes of this work with

observations from the field will further provide insights into the importance of size-selective harvesting as a driver of life history evolution in wild populations.

The intensity of the selection here imposed is strong, but not unrealistic for some semelparous harvested species like salmon where up to 80-90% of individuals can get caught (Ricker et al. 1978; Hard et al. 2008). Fig. 1.5 gives an example of how artificial, knife-edge, selection may differ from that by fisheries on wild populations. The absence of a gradual size range over which fish are more or less vulnerable may have a considerable effect on the intensity of selection and is unrealistic for natural populations. However, such intensity is required to observe a response within the timeframe of an experiment. Furthermore, results can be translated to be relevant for more natural populations, as done by (Brown et al. 2008) for selection experiments on Atlantic silverside (Conover and Munch 2002). Our aim is to contribute to the understanding and identification of the evolutionary processes and



**Fig. 1.5.** Comparison between artificial selection regime on Atlantic Silversides (Conover & Munch, black solid line) and annual exploitation rates for Southern Gulf of Saint Lawrence cod (black broken line), Gulf of Maine cod (grey solid line), and Bering Sea Pacific cod (black dotted line). Length is plotted relative to a “standardized” length corresponding roughly to 50% vulnerability to the fishing gear and here demonstrates the difference in vulnerability to selection between experimental and natural populations. From: Hilborn and Minte-Vera 2008.

genes involved in a rapid response to selection. To our knowledge, this work represents the first case where artificial selection experiments in fish have been combined with measurements of both neutral genetic diversity at microsatellite loci and putative candidate loci. We hope that these findings will aid the understanding of contemporary evolution and adaptation of natural populations in general.

### **1.5.3 Outline of the experiments and thesis**

This work comprises three main data chapters. Firstly, I describe the phenotypic response to selection experiments on wild-caught Trinidadian guppies and infer heritability of traits and rates of evolution from the data obtained (Chapter 2). Next, I reveal the genetic dynamics of microsatellite loci and quantify stochastic effects and drift between selection lines and over generations (Chapter 3). Finally, I explore the effects of selection at putative candidate genetic markers (Chapter 4)

2.

**PHENOTYPIC DIVERGENCE OF MALE**  
***P. RETICULATA* IN RESPONSE TO THREE**  
**GENERATIONS OF ARTIFICIAL SELECTION**  
**ON MALE STANDARD LENGTH**

## **2.1 INTRODUCTION**

### **2.1.1 Phenotypic change and micro-evolution in natural populations**

The fact that evolution can be a rapid process, taking place over brief time periods and few generations, has become more widely acknowledged and accepted in recent years (Endler 1986; Hendry and Kinnison 1999; 2001; Kingsolver et al. 2001). Several different lines of evidence support the potentially rapid pace of evolutionary change. Firstly, long-term monitoring studies (Grant and Grant 2002; Avilés et al. 2006; Gratten et al. 2008) and manipulations (Endler 1980; Reznick et al. 1990) of wild populations are becoming more common and provide increasing support for natural selection affecting trait distributions on contemporary, ecological, timescales. Secondly, instances of rapid divergence and parallel evolution between populations only recently isolated from each other provide information on the pace adaptation can take in contrasting environments (Stearns 1983; Hendry 2000; Bell et al. 2004). Thirdly, common garden experiments and selection experiments have convincingly demonstrated the genetic basis of, and presence of genetic variation for, fitness-related traits and the capacity hereof to respond rapidly to intense selection (Conover and Present 1990; Houle-Leroy et al. 2003; Barrett et al. 2010). Finally, the increasing influence of human activities on natural populations provides a unique opportunity to study adaptation to environmental change in wild populations, and provides some of the strongest cases demonstrating rapid evolution (Kettlewell 1973; Heap 1997; Lee 2002).

Whether human activities increase the pace of evolutionary change in the wild, or whether they merely direct our attention more towards mapping such changes, in order to understand the magnitude of our impact on the natural environment, is as yet undecided (Carroll et al. 2007). The fact is that the ever-increasing demand for resources by man has affected the majority of ecosystems and species globally (Palumbi 2001; Stockwell et al. 2003; Darimont et al. 2009). Of particular interest here, from both a conservation- and a management perspective (Fenberg and Roy 2008), is the impact of size-selective fishing on exploited marine fish species. Phenotypic shifts in a range of (life history) traits have been reported for several fish

species that have been exposed to intense harvesting pressures (Jorgensen et al. 2007; Sharpe and Hendry 2009) and evolutionary change has been suggested as an explanatory underlying process. As discussed in 1.2.5, the majority of evidence for fisheries-induced evolutionary change (FIE) is based on phenotypic data and indirect measurements of evolutionary change, and care should be taken when interpreting phenotypic change as genetic change. For example, in their analysis of studies on climate change adaptation in birds, Gienapp et al. (2008) found that in only 40% of studies reporting selection on heritable traits, selection actually resulted in a measurable response in the expected direction. Furthermore, from the studies reporting both phenotypic and (quantitative) genetic trends associated with climate change, a majority (64%) of the phenotypic trends were not consistent with the genetic trends reported (Gienapp et al. 2008). Similarly, Hilborn and Minte-Vera (2008) found no correlation between the intensity of fishing and the magnitude of change in weight gain observed in their meta-analysis of 73 exploited stocks. Furthermore, they regularly observed increases in mass at age in exploited stocks as well; representing a response to selection in the opposite direction of what one would theoretically expect. Whether this represents the difficulties involved in detecting genetic change of highly plastic traits under weak selection, biased trait estimates due to the challenges involved in measuring quantitative traits in the wild, or a lack of genetic support for phenotypic trends, remains largely unresolved.

### **2.1.2 The required evidence for quantifying fisheries-induced evolution**

Although the range of studies suggesting FIE is impressive, conclusions on the magnitude and importance of FIE have been criticised on the grounds that (1) phenotypic evidence alone cannot be regarded as conclusive evidence for evolutionary change and (2) careful assessment of the importance of FIE to fisheries management should be made, rather than making premature claims on its impact on yield and the health of stocks (Hilborn 2006; Kuparinen and Merilä 2007; Law 2007; Hilborn and Minte-Vera 2008).

In order to respond to such criticisms, several key pieces of information are required: firstly, in order for the rate of fisheries-induced change to be quantified, reliable



estimates of fisheries selection differentials and heritabilities (Box. 2.1) in harvested populations have to be compared with observed phenotypic shifts. Given the intensity of fishing mortality and the convincing indications available for considerable heritability of body size (Stokes and Law 2000), significant differentials can be expected. However, exact estimates are still rare and hard to obtain: estimating fisheries selection differentials requires information on catch data as well as on that of the remaining individuals, in addition to an opportunity to disentangle fisheries from other selective pressures in the wild. Estimates have been made for body size in cod (Swain et al. 2007) and pike (Edeline et al. 2007) and growth in several species of whitefish (Nusslé et al. 2011) and indeed suggest significant differentials imposed by fishing. Direct estimates of heritability of traits are still harder to obtain in the wild since they require trait values from both offspring and parents, something which is unfeasible for most marine fish species. Successful attempts by Funk et al. (2005) & Smoker et al. (2000) on salmon however, indicate a heritable component to body size of around 0.3, which is in correspondence with observations from experiments and other species (Law 2000; Conover and Munch 2002).

In addition to verifying any correspondence between fisheries selection pressure and observed phenotypic shifts, the molecular genetic basis of this response and of the traits concerned, like body size and the maturation process, needs to be better understood. Only empirically-derived genetic data can provide the conclusive evidence to demonstrate whether current and past selection by size-selective harvesting can induce genetic change over short periods of time. As explained in 1.2.7, such data are still rare and hard to obtain in the wild, although considerable effort is being made (Nielsen et al. 2009b; Jakobsdóttir et al. 2011). Currently however, studies on exploited fish species combining quantitative and molecular genetic data with phenotypic shifts and likely estimates of fisheries selection differentials are absent.

### **2.1.3 Artificial selection experiments for studying fisheries-induced evolution**

In light of the above, artificial selection experiments are an important tool for disentangling genetic and environmental components of phenotypic change (Brakefield 2003; Fuller et al. 2005; Conover and Baumann 2009). By minimising environmental variation, a genetic basis to traits of interest can be verified and estimated. The nature of the response can be precisely quantified and correlations between traits investigated (Chippindale et al. 1997; Houle-Leroy et al. 2003). Careful design of experiments allows for replication and the investigation of the generality of the response (Rose et al. 2005). Comparisons of experiments to wild populations facilitate valuable inferences on the selective forces shaping evolution in the wild (Korsten et al. 2010; Rand et al. 2010). Despite limitations of laboratory-based experiments, often caused by practical restrictions to the number of generations investigated, population sizes and the limited species suitable for such manipulations, these advantages make selection experiments extremely valuable for evolutionary studies. For the study of FIE in particular, experiments are of great value as they can combine the three different components to a holistic assessment of FIE by: (1) knowledge of- and control over selection differentials, (2) the possibility to accurately estimate heritability and quantitative genetic parameters and (3) the opportunity to monitor genetic change over generations at loci of interest in parallel with phenotypic shifts.

Here we present the response to artificial selection on standard length in replicate selection lines of the Trinidadian guppy, *Poecilia reticulata* L. descended from a wild-caught guppy population, as a model to study evolutionary effects of fishing. To our knowledge, this is the first time that selection on body size, as so well documented to be a major driver of differentiation of wild guppy populations (see 1.4), has been imitated under controlled conditions for this species. The results of the experiments conducted here therefore provide a valuable contribution to our understanding of the strength of selection by predation in wild guppies and fish species in general, and to our understanding of the complexity of selective forces that shape life history evolution in the wild. Trinidadian guppies have a unique niche as one of the best-known examples of life history evolution in the wild (Magurran 2005). Due to the

intrinsic qualities of the species as a model species (see 1.4.1) and its capacity to respond quickly, and in a predictable fashion, to selection in the wild, we can investigate the underlying genetic mechanisms in more detail to elucidate the rate of evolution, genetic basis and heritability of body size in these extraordinary little fish.

In this chapter, the phenotypic response to selection on body size is addressed. We use the results of our experiments to make inferences on heritability of this ecologically important trait and provide estimates of quantitative genetic variation and rates of evolution. The molecular genetic response to selection in our experiments will be discussed in Chapters 3 and 4.

## **BOX 2.1. TERMINOLOGY IN QUANTITATIVE GENETICS**

### ADDITIVE GENETIC VARIANCE

The proportion of the genetic variation that is derived from paternally and maternally inherited genes: the heritable amount of genetic variation in the narrow sense.

### RESIDUAL VARIANCE

The non-additive component of phenotypic variation, comprised of environmental variance and non-additive genetic variance

### NON-ADDITIVE GENETIC VARIANCE

Dominance: Interaction between alleles at the same locus or between loci (epistasy) where one allele conceals the expression of another

Pleiotropy: Phenomenon where one gene affects multiple traits

Linkage Disequilibrium: The correlation of allele frequencies at two (or more) polymorphic loci

### HERITABILITY ( $h^2$ )

Narrow sense: The proportion of phenotypic variance present for a trait which is attributable to additive genetic variance. In other words: the proportion of the phenotype determined by the genes inherited from the parents, or the degree of resemblance between related individuals.

Broad sense: The proportion of variance present for a trait which is genetically determined, comprising both additive and non-additive genetic variance.

### SELECTION DIFFERENTIAL ( $S$ )

The difference between the mean phenotypic value of the population as a whole and the mean of the selected parents. It is equivalent to the selection intensity  $i$  multiplied by the phenotypic standard deviation  $\sigma_p$ .

### SELECTION RESPONSE ( $R$ )

The change in mean phenotypic value between the mean phenotypic value of the population as a whole and the mean of the offspring. It depends on both heritability ( $h^2$ ) and the selection differential ( $S$ ).

### BREEDER'S EQUATION

The relationship between the response to selection  $R$ , the selection differential  $S$ , and the narrow-sense heritability  $h^2$ :

$$R = h^2 S$$

### ADAPTIVE GENETIC VARIANCE

Heritable phenotypic variation that is influenced predominantly by natural selection, so enhancing fitness in specific environments (Carvalho et al. 2003).

Unless stated otherwise, definitions are obtained from Falconer and Mackay (1996)

## **2.2 EXPERIMENTAL DESIGN**

### **2.2.1 Rationale behind chosen selection regime**

For the purpose of this study, a selection regime was required that was: (1) straightforward to execute and manipulate; (2) constant for all generations and insensitive to potential slight variations in generation time and to expected differences in body size, growth rate and maturation between lines; (3) controllable for- and excluding as much as possible – complicating factors like maternal effects and reproductive variance; (4) severe enough to likely give a response within three generations of selection and (5) maintainable by one person for multiple generations of fish within a single aquarium system. For these reasons, it was decided to perform mass selection experiments in which selected individuals were allowed to breed freely amongst themselves for a fixed amount of time, after which a cut-off selection regime based on adult (maximum) male standard length (SL) was used to select individuals for breeding subsequent generations.

### **2.2.2 Minimising effects of drift and sexual selection**

The choice of population used in our experiments was motivated by three main points: (1) ensure high levels of genetic variation in the baseline population; (2) minimise reduction of neutral genetic variation over generations and (3) ensure sufficient offspring to facilitate short generation times and minimal age-related variation in body size.

In order to maximise genetic variation in the starting population, fish from a single high predation population were used as the starting material for selection, as high predation populations have naturally higher levels of genetic variation (Shaw et al. 1994; Suk and Neff 2009). Furthermore, female guppies are capable of storing sperm and broods from high predation populations have high levels of multiple paternity compared to low predation populations (Kelly et al. 1999). Sampling of gravid females therefore ensured greater variation in the F1 generation than the wild-caught sample.

In order to decrease reproductive skew resulting from female choice and sexual selection, and to maximise the likelihood of a large number of males fathering in each generation, male competition was reduced by having a female-biased sex ratio of 1.5. Houde (1997) observed that naïve, virgin females are more responsive and less discriminating towards approaching males than experienced females. Rearing the sexes separately in our experiments may thus result in less pronounced effects of sexual selection on reproductive variance than expected under more natural rearing conditions. Furthermore, males from high predation localities perform higher frequencies of sneaky matings (Matthews et al. 1997) than their low predation counterparts and the use of a high predation population therefore may have reduced female choice even further.

The Lower Tacarigua was ultimately chosen for its abundance of fish, thereby not likely to be negatively affected by our sampling of individuals, the easy access for sampling and its prolific breeding properties (out of four sampled populations it produced easily the most offspring).

### **2.2.3 Minimising maternal effects and inbreeding**

In order to remove maternal effects (Box 2.2) and to ensure effects of environmental variation on body size were minimised and standardised for all fish, two generations of random breeding were allowed prior to selection. For selection, a cut-off selection regime was used with an intensity of 20%. This percentage was chosen as we considered it sufficient to provoke a response to selection within the planned number of generations but not as severe as used by others, which has been criticised for being unrealistically high and knife-edged; selecting every individual above a single cut-off size (Brown et al. 2008; Hilborn and Minte-Vera 2008). Although in natural systems, selection is not likely to be as knife-edge as imposed here, such high levels of exploitation mortality do occur in harvested populations (Mertz and Myers 1998, Fig. 1.4).

The decision to use random females and only select on adult male body size was driven by the need to minimise inbreeding levels and variation in maternal effects between lines. Using differently sized females for breeding in the different lines could increase the impact of maternal effects, as female body size is known to affect offspring quality (Reznick et al. 2001b). Secondly, as males virtually stop growth after maturation (Reznick 1990) selection on adult male body size can ensure individuals are chosen based on their genetically-determined body size. Female guppies on the other hand continue growth throughout their adult life and selecting females based on their size would only ensure same-aged, and therefore more likely to be related, individuals to be selected. This could increase inbreeding within selection lines and is therefore undesirable.

### **BOX 2.2. MATERNAL EFFECTS**

Maternal effects are a specific source of environmental variation that affects a female's offspring, causing resemblance between offspring and their mother or between offspring of the same mother, and thereby can be mistaken for inherited genetic variance (Falconer and Mackay 1996). Most maternal effects are pre- and post-natal nutritional effects on offspring depending on the mother's body condition. For example, larger mice providing more milk to their young, increasing weaning weight (Young et al. 1965) and older female guppies give birth to smaller fry (Reznick et al. 2001b). Alternatively, maternal effects can arise from the fact that a mother's offspring share a common environment, resulting in resemblance between offspring traits but not necessarily depending on the condition of the mother. Effects of egg incubation temperature on duckling size (Hepp et al. 2006) provide an example.

Maternal effects can bias experimental results considerably (Falconer 1953) and if left unaccounted for can be of great concern in interpreting experimental results and be of effect even over two generations (so grandmaternal effects on grandoffspring, Kirkpatrick and Lande 1989). In many cases however, (except when one performs selection on maternal effects themselves for example) maternal effects can be standardised between experimental groups by rearing individuals for one or two generations in a common environment.

## **2.3 MATERIAL AND METHODS**

### **2.3.1 Sampling and transport of fish to the Bangor University aquarium facility**

Ninety male and 90 female guppies were collected from the Lower Tacarigua River (10°38'49.5"N, 61°22'47.2"W, grid reference 677228E/117420N). Fish were caught using a 1m<sup>2</sup> fine meshed net (4mm<sup>2</sup>) net and hand nets, maintained alive in 20L buckets and transported to the aquarium facility at the University of the West Indies in Tunapuna, where male and female fish were kept in separate 50L aquaria for 48 hours. Fish were transported to the UK in groups of 9-10 fish per 1.5L bag, filled with 0.5L of water and pure oxygen. Upon arrival at Bangor University, males and females were divided equally between two 140L (122 x 30.5 x 38.1 cm) tanks with gravel substrate and artificial vegetation. Only 2 fish died during transport and all fish were in the new aquarium facility within 48hrs.

### **2.3.2 Health monitoring**

Throughout the experiment water quality and health of fish were monitored closely. Unhealthy looking fish were immediately removed from the stock tanks and treated with methylene blue or antibiotics. If no improvement was observed over a seven-day period, the individual was killed humanely using an overdose of MS222 (tricaine methane sulphonate, 2.5ml/L) followed by destruction of the brain. Dead males were placed in 100% ethanol and stored at room temperature. If improvement was seen, the fish was kept in isolation for a further seven days after symptoms had disappeared and then returned to its original tank.

### **2.3.3 Rearing protocol**

Fish were kept in a controlled temperature environment with 12:12 hour light cycles. Water temperatures ranged from 24.5-25.2°C and pH was 7-7.3. Fish were fed *ad libitum* live brine shrimp (*Artemia artemia*) nauplii in the afternoon, except during some weekends and Bank Holidays, when they were fed commercial flake food. Tanks were inspected for fry daily and fry were immediately transferred to separate tanks where they were reared in densities up to 2.08 fry/L. When the oldest fish had



reached 21 days of age and onwards, juveniles were checked for maturity status daily and maturing individuals were removed to an identical separate, single-sex 120L aquarium with densities up to 1.25 individuals/L. Males and females were separated well before reaching sexual maturity and kept in single-sex tanks until all had been sexually mature for at least 30 days, at which all fish was measured for standard length (SL). At the end of the breeding period in each generation except the F6, all parental fish were measured again. Of all male fish still alive at the end of the breeding period, tissue samples were taken in addition to size measurements, placed directly in 100% ethanol and stored at room temperature until DNA extraction.

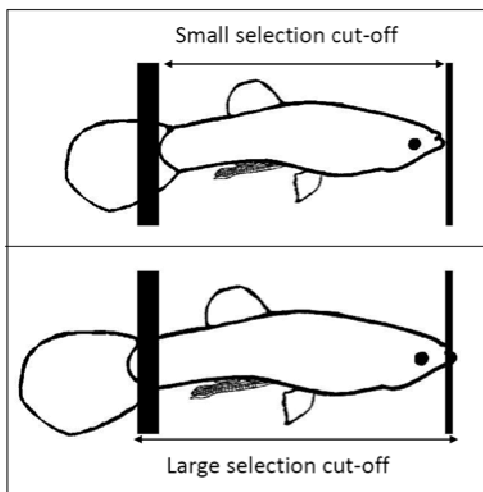
#### **2.3.4 Rearing of random breeding generations: removal of maternal effects and increasing population size in generations F1-F2**

For the first generation of breeding in the aquarium, fry (F1) were collected for 60 days in separate 105L (91.5 x 30.5 x 38.1 cm) aquaria and a total of 313 fry were collected and reared to maturity. From these, 50 males and 50 females were chosen at random and used to parent the F2 generation. In total, 1017 F2 fry were collected over a 90-day period. From these, a random fraction of 97 fry were taken to estimate maturation age and size (see further 2.3.7). A further 250 males and 375 females were reared to maturity, randomly divided into five groups of 50 males and 75 females, placed in five 140L aquaria and used to parent the F3 generation. F3 were produced over a 60-day period and 1200 individuals were reared to sexual maturity. These fish were used to set up the experimental selection lines S1, S2 (selection for small male body size) L1, L2 (selection for large male body size) and a random mating control line (C).

#### **2.3.5 Establishing selection lines in generation F3**

All fish used in selection lines were kept at constant 12:12hr light cycles in identical 120L aquaria (60 x 50 x 40 cm) over 2 levels in a continuous flow-through system with coral sand substrate and artificial vegetation. Breeding and rearing tanks from all treatments were distributed at random throughout the bank, water temperatures ranged from 24.5-25.2 °C and pH was constant at 7.2-7.5.

When all males had been mature for at least 30 days, 50 mature males were randomly chosen out of 550 mature F3 males, measured and designated to the random breeding control line (from now on referred to as C). The remaining 500 mature males were all measured and the 20% extreme percentiles of the obtained normal distribution were determined as the selection cut-off size. The 20% smallest and 20% largest male fish were selected subsequently by aligning individual fish to a chart on which the cut-off size was outlined (Fig. 2.1). When the number of selected fish did not match the required sample size of 50 individuals, previously discarded or selected fish (depending on whether too many or too few fish were selected initially) were rechecked until the required number of fish was obtained. In all cases, the difference between determined size and the number of fish initially selected was always less than nine individuals, even when size differences between fish were



**Fig. 2.1.** Representation of knife-edge selection procedure, using a chart with two bars. Males falling within the two bars were selected for breeding in S-lines, fishing falling outside both bars for breeding in L-lines. All other fish were discarded.

around 0.1 mm, which strengthens confidence in the accuracy of our measurements.

The selected small and large fish were each randomly divided over two groups, resulting in four groups of 50 males each, two of which consisting out of the smallest 20% of males (from now on referred to as S1 and S2) and two out of the largest 20% of males (from now on referred to as L1 and L2). Each group of males was partnered with a group of 75 randomly chosen F3 females to parent the F4 generation.

### **2.3.6 Maintenance of selection lines in generation F4-F6**

Selected F3 individuals were left to breed at will for 30 days and any fry observed during this period discarded. After this period, F4 fry were collected daily for 60 days. At irregular intervals but at least once weekly, fry were measured immediately after collection until between 100 and 150 fry per line were measured for size at birth. Fry

were kept in identical 120L aquaria and the F4 was reared as described previously, in separate aquaria for males and females until all males were mature for at least 30 days.

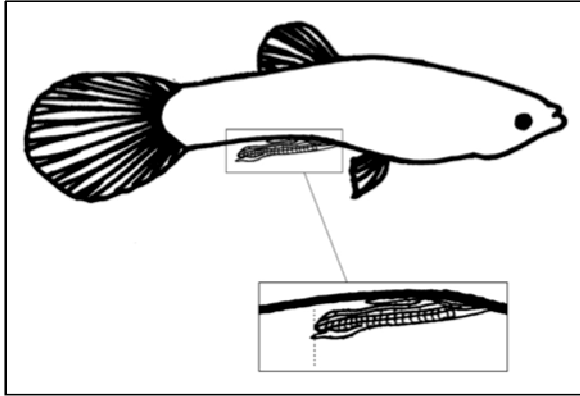
When all males had been mature for at least 30 days, 75 F4 females were chosen at random from each line and measured. Per line, 250 F4 males were selected at random and photographed in the same manner. For these male fish, the size distributions obtained were used to determine the selection cut-off size, which was decided upon such that only the smallest/ largest 20% of males (50 fish) were retained. For the control line, 50 fish were selected at random from the pool of 250 males and measured again.

All selected F4 fish were left to mate freely for 30 days to produce the F5, which were reared and of which 50 mature males and 75 females per line were selected to produce the F6 in the same way. For the F6, a random fraction of 100 fry per line was taken to estimate maturation age and size (see 2.3.7) and approximately 100 males and 100 females were reared in the main experiment as for previous generations. When all F6 males were mature for at least 30 days, all fish were measured and the experiments terminated. A schematic overview of the experiments is provided in Fig. 2.2.

### **2.3.7 Estimation of maturation age and size in the F2 and F6 generations**

In the F2 and F6 fry, a random fraction of fry was captured daily and placed in 4 L plastic sweet jars. Up to six fry born on the same day were placed in a jar. All jars were aerated and kept in a controlled temperature environment at 25°C. They were fed *ad libitum* brine shrimp daily and a full water change was performed once per week, at which point each jar was moved to a randomly chosen location within the controlled temperature room. Randomisation was done by collecting all jars on the floor and replacing them at will throughout the room, in order to randomise any effects of microclimate on maturation. Furthermore, in the F2 generation, SL of all fish was measured weekly by taking the average of three measurements using electronic callipers.

Maturity status was confirmed daily by visual inspection of all fish over 14 days of age. If a fish was female, recognisable by black pigment speckling in the anal area, she was removed from the experiment. If a male, recognisable by the onset of gonopodium development, the fish was kept in the jar and reared to maturity. A maximum of five males were reared in any one jar. Upon reaching sexual maturity,

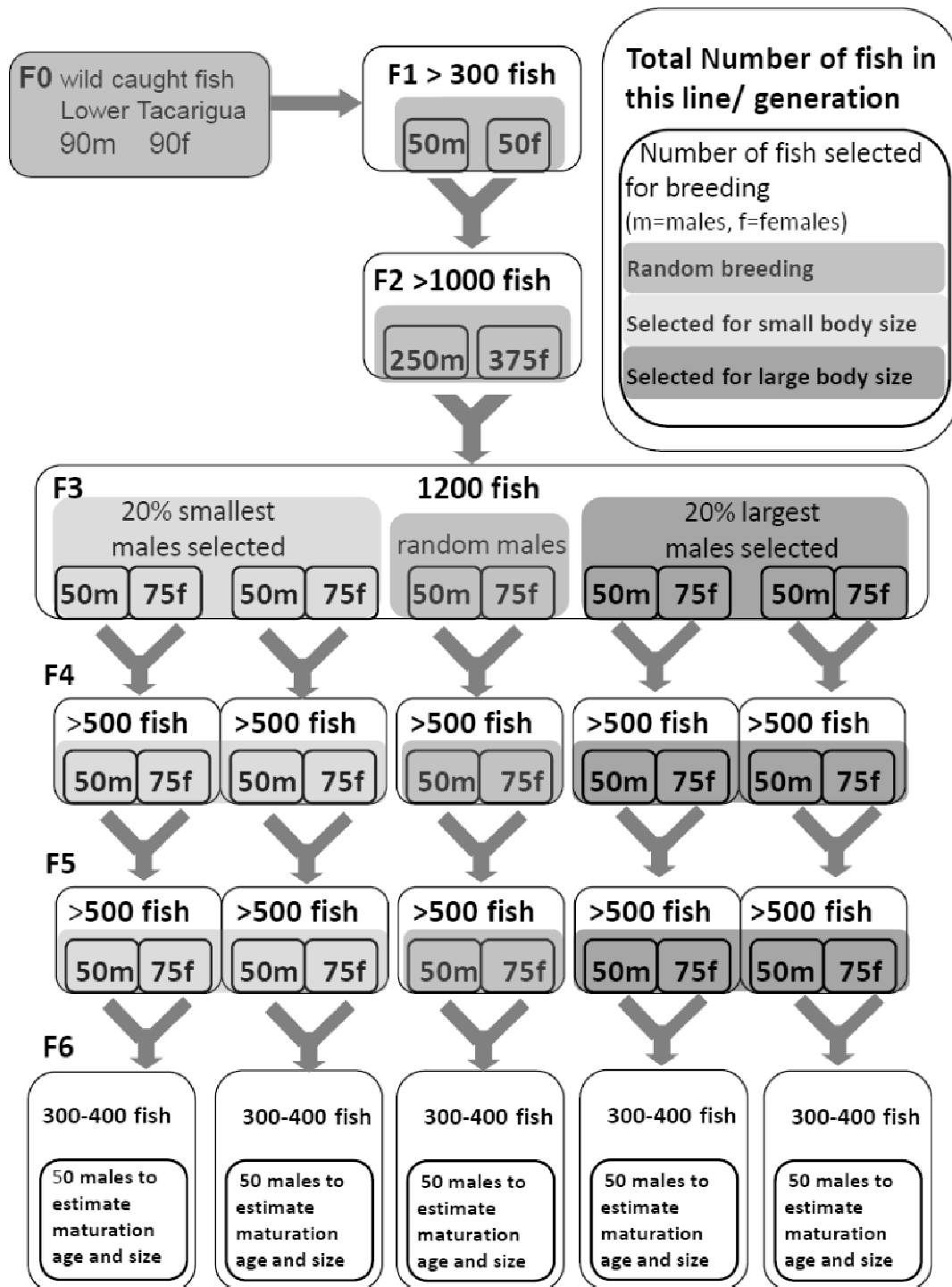


**Fig. 2.3.** Mature male guppy, as indicated by the developmental stage of the gonopodium: maturation is defined as the fleshy hood extending beyond the tip of the gonopodium (inset).

as indicated by the fleshy hood extending beyond the gonopodium tip (Fig. 2.3, Houde 1997), the male was photographed and in the F6 generation a tissue sample was taken after which he was removed from the experiment. In the F2, males were kept in jars for a further 30 days to obtain a complete representation of male growth in relation to maturity status.

### **2.3.8 Measurements**

Fish were measured by photographing them in a small Petri dish with mm scale on a lighted surface. Pictures were analysed using Adobe Photoshop® to obtain standard length of all individual fish. Accuracy and repeatability of measurements was verified using a set of 50 males for which three pictures were taken for each male and in addition three measurements of standard length were made using electronic callipers. Repeatability was calculated as:  $\sum_{1-50} (1 - (\sigma_k^2 / \sigma_t^2)) / 50$ , in which  $\sigma_k$  is the standard deviation between repeated measures of the  $k^{\text{th}}$  fish and  $\sigma_t$  the standard deviation over all measurements (Lynch and Walsh 1998). Repeatability among measurements was high; 0.98 ( $\pm 0.026$  st. dev.) for caliper measurements and 0.95 ( $\pm 0.050$  st. dev.) for photographic measurements. Despite their slightly lower repeatability, photographs were used for the majority of the measurement in this study because of the flexibility allowed in the timing of analysis of the photographs and secondly, because the pictures can be stored and thereby it was possible to recheck them for possible erroneous measurements at a later time.



**Fig. 2.2.** Schematic representation of selection experiments, showing the number of fish reared in the different generations F0-F6. Numbers of fish selected and used for breeding the next generation are indicated in grey, the different shades indicating the different treatments: light grey for small-selected lines, dark grey for large-selected lines and intermediate shading for random breeding generations.

### **2.3.9 Statistical analyses**

PASW Statistics 18.0.3 for Mac was used to test for differences in standard length, age and size of maturation between the selection lines and control line, using single-factor nested analysis of variance (ANOVA), with treatment as fixed factor and replicate line as random factor nested within treatment. Bonferroni post hoc tests were used to reveal which comparisons were significant. Tests were done on both male and female traits and for each generation independently.

Linear regression was used to investigate changes in standard length over generations, except in cases where only two values were compared. In these cases, permutation tests were performed, using Rndom Pro 3.14 (Jadwiszczak 2009) with 10000 permutations.

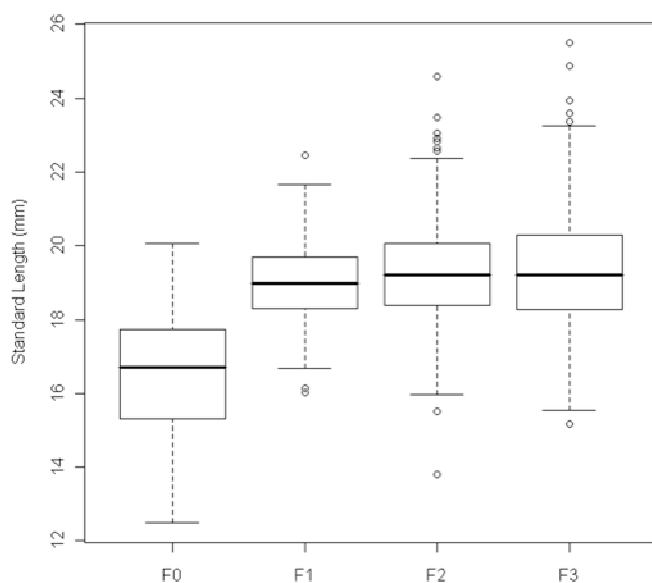
Quantitative genetic parameters were estimated following Falconer and Mackay (1996), (see Box. 2.1) using the standard relationships  $V_A = h^2V_P$  and  $R = h^2S$ . The selection differential  $S$  was calculated for each generation as the difference between mean SL of the population before selection, and mean SL of the selected parents. Response to selection  $R$  was calculated as the difference between mean SL of the population before selection and mean SL in the next generation. Heritability was estimated from the slope of the linear regression of the cumulative response to selection over generations;  $R/S$ . In order to correct for the fact that selection was performed on a single sex only, which implies the relation between heritability, selection differential and response to selection is;  $h^2=2RS^{-1}$ , heritability equals twice the slope of this regression. Note that this will inflate the error variance and could lead to an overestimation of actual heritability if some of the response to selection is mediated to Y-linked genes (Winge 1922b; Houde 1992; van Oosterhout et al. 2003).

## **2.4 RESULTS**

### **2.4.1 Breeding efficacy**

The first F1 fry were observed immediately after transport of wild-caught fish and breeding commenced directly. Throughout all generations, fry were typically found at least 30 days after breeding commenced, confirming gestation times found elsewhere (Evans and Magurran 2000).

The number of fry observed varied considerably between days, lines and generations. The maximum range difference in mean number of fry observed per line within a single generation was 15.4 / day in the F5 (the most prolific generation) and generation means ranged from 4.4 – 31.5 fry collected per day over all lines and generations. While reproductive output of females may have varied between lines and generations, different levels of cannibalism are also likely to have caused the variation, as fry were collected only once daily and females were regularly observed chasing newborns as soon as they left their hiding places. The main aquarium facility



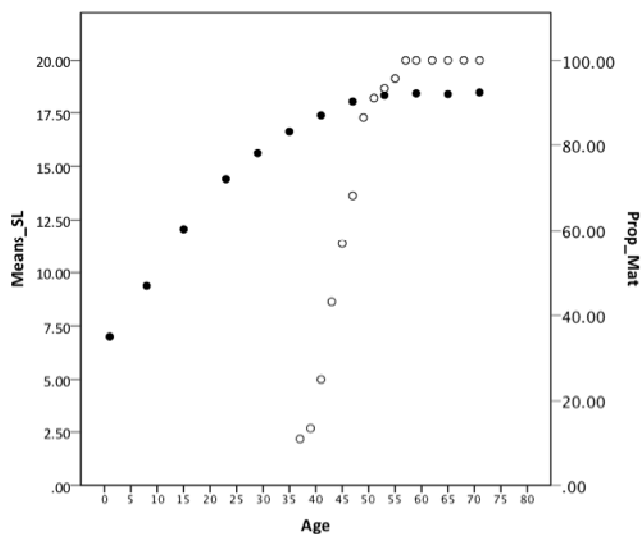
**Fig.2.4.** Box plot of standard length in mm of wild-caught males (F0) and the first three generations produced by random breeding (F1-F3). Differences in standard length are significant only between F0 and F1 males ( $p = 0.001$ ). Sample sizes were  $N = 84, 49, 252$  and  $542$  individuals respectively for the subsequent generations.

where the fish were housed contained multiple other experimental aquaria and disturbance by human presence cannot be excluded from having caused stress in the fish, increasing cannibalism amongst females. As reproductive output as such has not been quantified for individual females here, we refrained from making any further conclusions on the reasons behind the observed variation in offspring number.

## 2.4.2 Experimental results – random breeding generations

### Standard length

Rearing fish in the aquarium yielded larger fish than those caught in the wild (Fig. 2.4). The first generation of breeding in the lab yielded significantly larger males than those observed in the wild-caught sample (two-sample randomisation test with 10000 randomisations,  $p = 0.0001$ ). Two more generations of random breeding in the lab did however not result in changes in standard length of male fish ( $p = 0.122$  and  $0.647$  for the F1-F2 and F2-F3 comparisons respectively). Linear regression on standard length in generations F1-F3 confirmed this ( $R^2 = 0.003$ ,  $F_{1,841} = 2.289$ ,  $p = 0.131$ ). These results provide strong support for the assumption that optimal rearing conditions and *ad libitum* feeding in the aquarium environment resulted in maximal growth of fish and minimised the effects of variation in rearing environment and feeding levels on standard length.



**Fig. 2.5:** Growth and maturation in the F2 generation. Closed circles show mean standard length (mm) and open circles the proportion of males mature at a certain age (in days). Females were removed from the experiment as soon as they could be identified (14-21 days of age) but data points for juveniles include both male and female measurements. Growth follows a logarithmic model:  $(SL = 3.0963 * \ln(\text{age}) + 5.301)$ .  $R^2 = 0.920$ ,  $F_{1,10} = 115.188$   $p = 0.000$

### Growth rate

Growth rate was estimated in the F2 generation and showed a logarithmic growth pattern: ( $R^2 = 0.920$ ,  $F_{1,10} = 115.188$ ,  $p = 0.000$ ). Between 47 and 53 days of age, growth virtually stopped and all fish were mature at 57 days of age (Fig. 2.5). These data strongly support our assumption that the selection cut-off size in our experiments, which is defined when all males have been mature for 30 days, was determined on fully grown males and no males were selected that had not yet attained their maximum size.



### **2.4.3 Experimental results – response to selection**

#### **Response to selection in male standard length**

Twenty percent selection on male standard length resulted in strong divergence in length between treatments, but not between replicate selection lines within treatments (Fig. 2.6). A significant divergence in SL was observed between treatments after one generation of selection (nested ANOVA,  $F_{2,2.092} = 87.016$ ,  $p = 0.010$ ). A Bonferroni post hoc test revealed all treatments were significantly different from each other ( $p = 0.000$  for S-L and C-L and  $p = 0.023$  for S-C). Replicate lines within treatments did not differ significantly from each other ( $F_{2,1.211} = 1.041$ ,  $p = 0.354$ ). After two generations of selection (in F5 males), this pattern was repeated ( $F_{2,1985} = 412.277$ ,  $p = 0.003$  between treatments,  $F_{2,1.230} = 1.037$ ,  $p = 0.335$  within treatments). For the F6, the variance in sample size between lines was greater than in previous generations (Table 2.1) and the design was therefore unbalanced. Creating a balanced design by sampling 68 individuals per line at random demonstrated significant differences between treatments ( $F_{2,2.800} = 14307.362$ ,  $p = 0.000$ ) but not within treatments between lines ( $F_{2,336} = 0.008$ ,  $p = 0.992$ ).

**Table 2.2.** Summary of linear regression analysis for males in individual breeding lines of the F3-F6 generations. The slope of the regression is shown, as well as the  $R^2$ , F coefficient and  $p$  values.

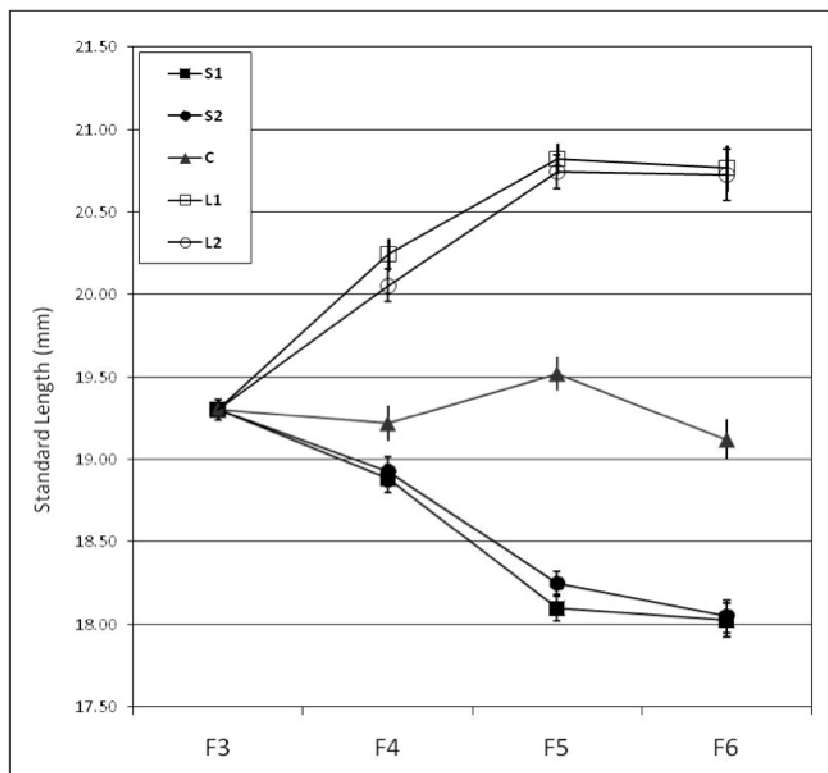
	<b>S1</b>	<b>S2</b>	<b>C</b>	<b>L1</b>	<b>L2</b>
<b>Slope</b>	-0.495	-0.463	0.007	0.616	0.614
<b>R<sup>2</sup></b>	0.13	0.113	0	0.151	0.133
<b>F</b>	$F_{1,1158} 172.948$	$F_{1,1144} 145.471$	$F_{1,1145} 0.030$	$F_{1,1142} 203.001$	$F_{1,1103} 169.456$
<b>p</b>	0	0	0.863	0	0

**Table 2.3.** Summary of linear regression analysis for females in individual breeding lines for the F3-F6 generations. The slope of the regression is shown, as well as the  $R^2$ , F coefficient and  $p$  values.

	<b>S1</b>	<b>S2</b>	<b>C</b>	<b>L1</b>	<b>L2</b>
<b>Slope</b>	0.686	0.514	0.704	0.643	0.653
<b>R<sup>2</sup></b>	0.099	0.058	0.102	0.083	0.087
<b>F</b>	$F_{1,594} 65.066$	$F_{1,598} 36.495$	$F_{1,592} 67.304$	$F_{1,595} 54.115$	$F_{1,595} 56.671$
<b>p</b>	0.000	0.000	0.000	0.000	0.000

**Table. 2.1.** Overview of standard length, standard deviation and population size of selection lines and control line in generations F3-F6. Due to time limitations, the experiment was halted in the F6, resulting in fewer samples and greater variation in population size than in previous generations.

		S1	S2	C	L1	L2
F3	SL (mm)	19.30				
	St. dev.	1.49				
	N	542				
F4	SL (mm)	18.88	18.92	19.22	20.24	20.06
	St. dev.	1.32	1.39	1.60	1.53	1.57
	N	244	245	232	248	247
F5	SL (mm)	18.09	18.25	19.52	20.82	20.74
	St. dev.	1.15	1.08	1.64	1.44	1.60
	N	244	245	249	249	249
F6	SL (mm)	18.02	18.05	19.12	20.77	20.73
	St. dev.	1.16	1.05	1.37	1.47	1.26
	N	130	114	124	105	68



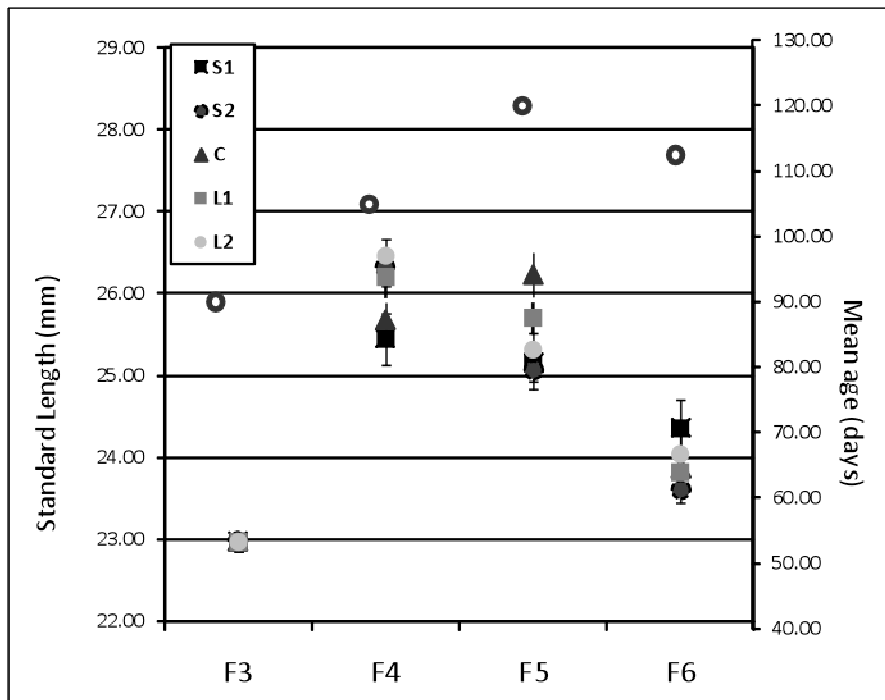
**Fig. 2.6.** Standard length in mm of selection lines and random breeding control line in generations F3-F6. Open circles and squares are replicate lines selected for large standard length, closed circles and squares for small standard length and the control line is given by closed triangles. Error bars indicate standard errors of the mean.

Linear regression revealed standard length declined significantly in both small selected lines and increased significantly in both large selected lines. No change in standard length was observed for the control line (Table 2.2).

### **Response to selection in females**

No response to selection was observed in females. Female SL was much more variable than male SL and varied between lines and generations (Fig. 2.7). No significant difference between treatments was observed for the F4 ( $F_{2,1.997} = 0.859$ ,  $p = 0.538$ ) but differences between lines within treatments were significant ( $F_{2,368} = 3.315$ ,  $p = 0.037$ ) indicating that environmental heterogeneity was greater than the effect of directional selection, and that the response to selection in males was largely the result of selecting genetic variation at Y-linked genes. For the F6, a similar conclusion can be drawn ( $F_{2,2.001} = 0.032$ ,  $p = 0.969$  between, and  $F_{2,366} = 3.639$ ,  $p = 0.027$  within treatments). For the F5, significant differences between treatments were observed ( $F_{2,2.115} = 10.546$ ,  $p = 0.079$ ), which, a Bonferroni post hoc analysis revealed, were caused by the control being larger than both S-lines and L-lines ( $p = 0.000$  and  $p = 0.027$  respectively).

Although regression analysis did indicate a significant increase in SL of females over generations (Table 2.3), Fig 2.7 shows that this is not in a unidirectional pattern, nor consistent with the selection regime and therefore is unlikely to be a result of selection on standard length in males. As no trend was observed over generations and since the few significant differences observed were not in agreement with the direction of imposed selection, changes in female standard length are not likely to be the result of directional selection, nor likely to have had a reinforcing impact on the response to selection in males by consequential maternal effects. The variation in female size is most likely the result of stochastic variation in age distributions of the randomly chosen females. Female guppies have indeterminate growth and a significant correlation was observed between standard length and mean age of females at the time of measurement over generations (Pearson's correlation = 0.315,  $p = 0.000$ , Fig. 2.7).



**Fig. 2.7.** Standard length of females in generations F3-F6 for selection lines and control line plotted with mean age of females per generation (large open circles). Grey circles and squares represent replicate lines selected for large standard length, dark circles and squares for small standard length and the control line is given by closed triangles. Error bars represent standard errors of the mean.

### Size at birth

No significant difference in size at birth was observed between treatments in the F6 (nested ANOVA,  $F_{1,2.002} = 5544$ ,  $p = 0.143$  between treatments and  $F_{2,368} = 1.647$   $p = 0.194$  within treatments).

### Size and age of maturation

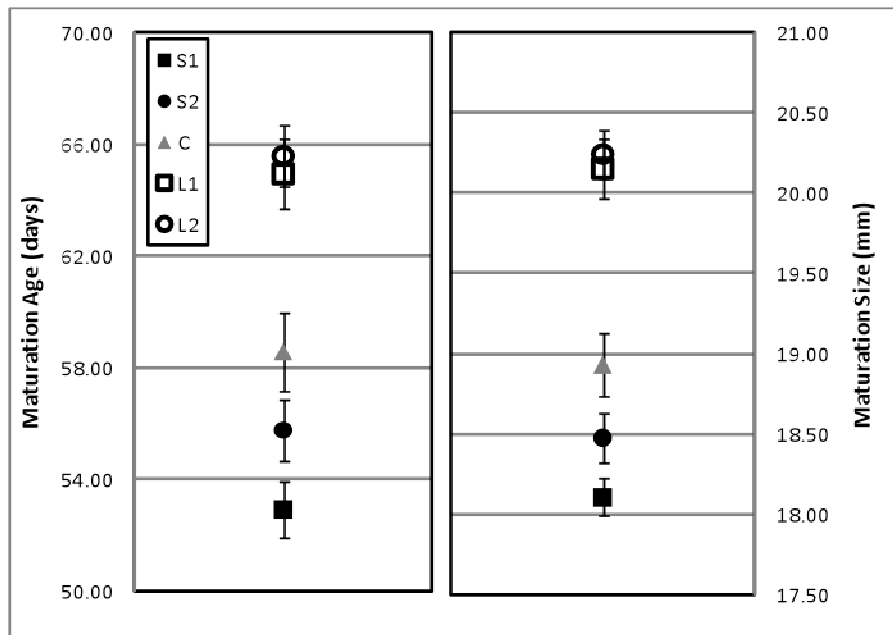
Size and age of maturation were estimated prior to setting up selection lines, in the F2 generation, and for each selection line independently in the final generation (F6).

In the F2, maturation age ( $Age_{mat}$ ) and standard length ( $Size_{mat}$ ) were 46.44 days ( $\pm 9.37$  st. dev) and 18.03 mm ( $\pm 1.03$  st. dev) respectively. This was significantly smaller than the mean standard length in the main F2 tanks (19.23 mm  $\pm 1.49$  st. dev). Even when allowing for an additional 0.5mm of growth after maturation, (as can be expected, see Fig. 2.5), the difference in standard length between fish used for the

maturation experiment and those in the main breeding tanks was still significant. However, complications were encountered with maintaining a constant room temperature where the maturation tanks were housed. Although no significant trend for reduced age or size of maturation over time was observed, ( $R^2 = 0.057$ ,  $F_{1,42} = 2.558$ ,  $p = 0.117$ ) room temperatures up to 2°C above the desired temperature of 25°C were regularly observed and cannot be excluded from having resulted in reduced maturation sizes, resulting in smaller fish than those observed in the main experiment.

In the F6 generation, such complications were absent and  $\text{Size}_{\text{mat}}$  was not significantly different from mean standard length in the control line, or from the F2. Furthermore, no significant difference was found between the F2 and F6\_C main breeding tanks.

Maturation age and size were highly correlated in the F2 ( $r = 0.651$ ,  $p = 0.000$ ) and the F6 ( $r = 0.799$ ,  $p = 0.000$  over all data and  $r = 0.495$ ,  $0.683$ ,  $0.795$ ,  $0.806$  and  $0.585$  for S1, S2, C, L1 and L2 respectively). Both  $\text{Age}_{\text{mat}}$  and  $\text{Size}_{\text{mat}}$  diverged significantly in the F6 generations between treatments ( $F_{2,1902} = 28.574$ ,  $p = 0.038$ , for  $\text{Age}_{\text{mat}}$  and  $F_{2,1895} = 51.284$ ,  $p = 0.022$  for  $\text{Size}_{\text{mat}}$ ) and both within-treatment comparisons were non-significant ( $F_{2,228} = 1.429$ ,  $p = 0.242$ , for  $\text{Age}_{\text{mat}}$  and  $F_{2,228} = 1.337$ ,  $p = 0.265$  for  $\text{Size}_{\text{mat}}$ ). The high  $p$ -value in these comparisons can be mainly explained by heterogeneity between lines in the S-treatment (Fig. 2.8). A Bonferroni post hoc test on a one-way ANOVA revealed that, whilst both L-lines were almost identical for both traits ( $p = 1.000$ ) and highly differentiated from both S-lines and the control ( $p < 0.001$  for all comparisons), S2 was not significantly different from the S1 ( $p = 1.000$  for both traits) and the control ( $p = 0.962$  for  $\text{Age}_{\text{mat}}$  and  $p = 0.460$  for  $\text{Size}_{\text{mat}}$ ).



**Fig. 2.8.** Size (SL in mm) and age (days) of maturation of selection lines and random breeding control line in the F6 generation. Open circles and squares are replicate lines selected for large standard length, closed circles and squares for small standard length and the control line is given by closed triangles. Error bars indicate standard errors of the mean.

#### **2.4.4 Estimates of heritability and quantitative genetic parameters**

##### **Heritability**

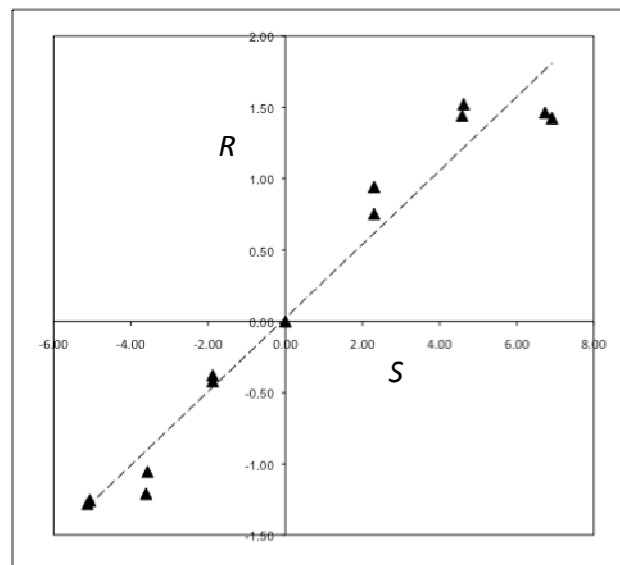
Realised heritability ( $h^2$ ) was estimated using the breeder's equation  $R=h^2S$  (Falconer and Mackay 1996) from the cumulative response to selection  $R$  over the cumulative selection differential  $S$ . Linear regression through  $R/S$  over all selection lines explained the majority of variation between data points (Fig. 2.9,  $R^2 = 0.960$ ,  $F_{1,14} = 335.230$ ,  $p = 0.000$ ) and thereby provides support for a heritable component to SL and for a symmetrical response to selection. As selection was done on a single sex only,  $h^2$  equals the slope of the regression multiplied by two, yielding in this particular case:  $h^2 = 2*0.258 = 0.518$ , with standard error estimated as twice the standard error of the slope;  $2*0.14 = 0.28$ . However, when estimated for small and large lines independently, the slope of the regression, i.e. heritability, did differ slightly between the different treatments and suggests a slightly larger  $h^2$  for small body size (Table 2.4). However this difference is not significant (independent sample

t-test, using variances computed from the standard error of  $h^2$  for  $n=7$ :  $T_{(2),11} = 1.31$ ,  $p = 0.218$ ). Although these two independent estimates did not differ significantly from each other, the low number of data points used for estimating  $h^2$  resulted in low power to test for statistical significance (effect size: 0.698, actual power: 0.225) and the lack of significance between these estimates is therefore more likely the result of a lack of power rather than the absence of actual significance.

Since reliable estimates of maturation age and size were obtained in one generation only, consequently the response to selection of these traits could not be measured, nor could heritabilities reliably be estimated. For these reasons, it was not possible to make inferences on genetic correlations between standard length, size and age of maturation.

**Table 2.4.** Estimates of heritability ( $h^2$ ) based on linear regression of  $R/S$  for small and large standard length independently. Slope of the regression,  $R^2$ ,  $F$ - and  $p$  value are given for each regression.

	S-Lines	L-lines
$R^2$	0.933	0.808
$F_{1,5}$	69.487	20.974
$p$	0.000	0.006
slope	0.269	0.199
S.E.	0.032	0.043
$h^2$	0.538	0.398
S.E. of $h^2$	0.064	0.086

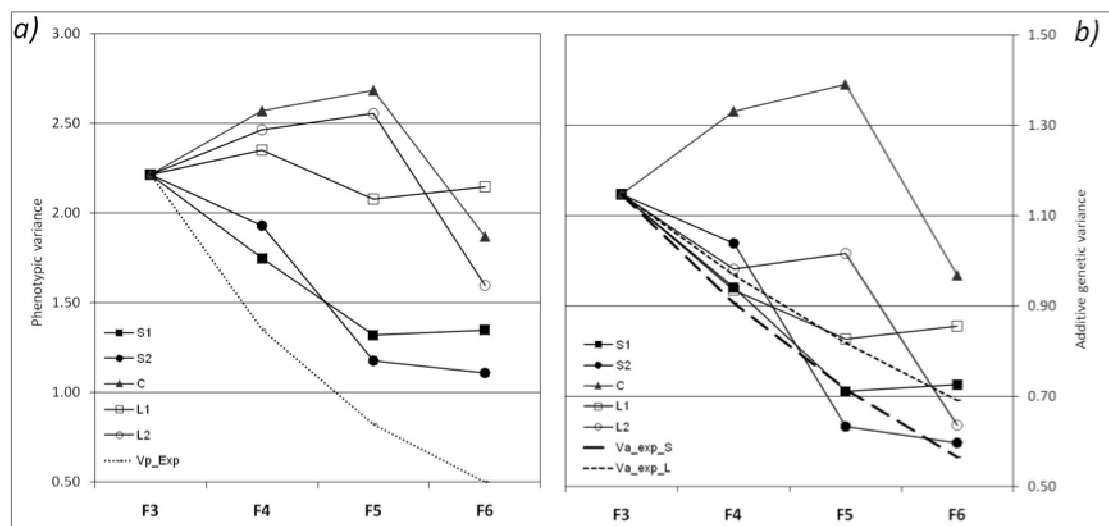


**Fig. 2.9.** Cumulative response to selection plotted against cumulative selection differential ( $S$ ) in mm. The linear regression through these data – dotted line – is an estimate of heritability, indicating symmetrical heritability of standard length. Slope = 0.258,  $R^2 = 0.960$ ,  $F_{1,14} = 335.230$ ,  $p = 0.000$ .

## Quantitative variance components of selection lines over generations

Using the estimates of heritability from Table 2.4, the additive genetic variance present in the different selection lines and generations can be estimated, following:  $V_A = h^2 V_P$ , where  $V_A$  is the level of additive genetic variation, and  $V_P$  the phenotypic variance, comprised of both genetic (additive and non-additive) and environmental variation (Falconer and MacKay 1996). Although heritability estimates for up- and down selection were not significantly different, the  $h^2$  estimates for S- and L-lines independently (Table 2.4) were used for the calculations of  $V_A$  for both directions of selection separately, as these estimates more closely resemble the actual changes in  $V_A$  than a global estimate. As no selection was performed on the random breeding control line, and therefore no R/S regression was available, the global estimate of  $h^2$  (0.518) from the regression was used to calculate  $V_A$  in this line.

Fig. 2.10 shows the phenotypic variation ( $V_P$ ) and its additive genetic component ( $V_A$ ) for all generations and breeding lines. Significant differences in  $V_P$  were found only between treatments in the F4 (ANOVA;  $F_{2,2} = 20.730, 14.450, 2.739$  and  $p = 0.046, 0.065$  and  $0.267$  for the F4, F5 and F6 respectively. Due to a lack of degrees of



**Fig 2.10.** Expected and observed phenotypic variance  $V_P$  (graph *a*) and its additive genetic component  $V_A$  (graph *b*) as estimated using:  $V_A = h^2 V_P$ , calculated for selection lines and random control independently. Expected  $V_P$  and  $V_A$  of selection lines were calculated according to Falconer and MacKay (1996) for 20% selection and  $i = 0.5 \cdot 1.4$ . Heritability estimates from Table 2.4 were used for S- and L- selected lines. For the F3 in the S- and L- lines and for F3-F6 in the control line, the overall estimate of  $h^2$  from the regression in Fig. 2.9 was used.



freedom, within-treatment comparisons and post hoc tests could not be performed).  $V_A$  did not differ significantly between treatments (nested ANOVA;  $F_{2,2} = 17.618$ , 16.436, 1.953 and  $p = 0.054$ , 0.057 and 0.339 for the F4, F5 and F6 respectively).

Furthermore, the two variance components did not portray the same trend of change over generations.  $V_p$  declined significantly over generations only in the S-Lines (logarithmic regression,  $R^2 = 0.930$ ,  $F_{1,5} = 32.183$ ,  $p = 0.002$ ) and did not change in either L-lines or control line ( $p > 0.149$  for all models).  $V_A$  did decline significantly in both directions of selection, (logarithmic regression,  $R^2 = 0.921$ ,  $F_{1,5} = 28.018$ ,  $p = 0.003$  for S-lines and  $R^2 = 0.820$ ,  $F_{1,5} = 10.268$ ,  $p = 0.024$  for L-lines), but not in the control line ( $p > 0.149$  for all models).

Using the observed values of  $V_p$  and the intensity of selection, the expected reduction in  $V_p$  and  $V_A$  can be calculated following:  $V_p^* = V_p (1-k)$  and:  $V_A^* = V_A (1-h^2k)$  where  $V_p^*$  and  $V_A^*$  are the respective variance components in the next generation and  $k$  is a constant depending on the selection intensity  $i$ . Here  $i = 1.4$  and  $k$  equals 0.7812, divided by two to account for selection on a single sex only, yields  $k = 0.3905$  (Falconer and MacKay 1996). Observed levels of  $V_p$  are higher than expected for both directions of selection but the difference is much more pronounced in L-lines than in S-Lines (Fig. 2.10). The observed reduction in  $V_A$  is very similar to that expected in S-lines and slightly higher than expected for L-lines.

#### **2.4.5 Rate of evolution**

Based on our results, the evolutionary rate of change in standard length was estimated in darwins and haldanes (Table 2.5). Both measures allow the comparison of the amount of phenotypic change in our experiments to measures of contemporary evolution elsewhere. The difference between the two measures lies in the parameter used for scaling evolutionary change. Darwins (d) measure change on an absolute timescale in millions of years. Haldanes depict evolutionary change as the change in standard deviation units per generation, thereby allowing for comparisons between species with contrasting life histories and generation times.

**Table 2.5.** Evolutionary rate based on three generations of selection on male standard length with  $i = 0.2$ .

Rate in darwins was calculated as  $(\ln X_2 - \ln X_1)/\Delta t$ , with  $X_2$  the mean standard length of the F6 and  $X_1$  of the F3. For darwins,  $t$  is originally calculated in  $10^6$  years. However, because of the temporal scale of our experiments,  $t$  here was calculated in years and evolutionary rates are shown as  $10^3$  Darwins (d). Rate in haldanes was calculated as:  $((X_2 / S_p) - (X_1 / S_p)) / g$ .  $S_p$  being the pooled standard deviation  $((n_1 - 1)S_1 + (n_2 - 1)S_2) / ((n_1 - 1) + (n_2 - 1))$  and  $g$  the number of generations (3). Formulas obtained from Hendry and Kinnison (1999).

	<b>S1</b>	<b>S2</b>	<b>C</b>	<b>L1</b>	<b>L2</b>
<b>10<sup>3</sup> Darwins</b>	-51.51	-50.51	7.11	54.94	53.51
<b>Haldanes</b>	-0.299	-0.296	-0.041	0.329	0.324

## **2.5 CONCLUSIONS AND DISCUSSION**

### **2.5.1 Response to selection**

Here we reported on selection experiments in Trinidadian guppies. Six non-overlapping generations of fish were reared in a common and constant environment and we demonstrated a rapid response to selection of male body size. Environmental variation in standard length was successfully standardised in the first two generations of breeding in the laboratory, indicated by an initial increase in standard length but no further change in size in generations F1-F3. After only three generations of selection, a decrease of 6.5-6.6 % and an increase of 7.4-7.6 % of standard length were observed in up- and down selected lines respectively. Laboratory rearing conditions resulted in an increase in standard length of fish. Over generations F3-F6 however, the random control line exhibited less than 1 % (0.94%) change. These results provide strong evidence for micro-evolutionary change in- and a heritable component to standard length in male guppies.

Furthermore, size and age of maturation diverged significantly between treatments. Relative to the age and size of maturation in the F6 generation of the random control, maturation age increased by 10.9 and 12% in L1 and L2 and maturation size by 6.4 and 6.9% respectively. In small selected lines, the response in maturation was slightly smaller for the S1 (-9.6% in maturation age and -4.4% in size) and much smaller in the S2 (-4.8 and -2.4% respectively). The greater response in large selected lines reflects the likewise greater response observed in standard length in L-lines, but the difference between both S-lines is not reflected in the response of standard length in these lines. It cannot be excluded that the rearing conditions in the individual rearing tanks of the maturation experiment affected the S2 fish in a way that resulted in increased standard length of these fish. This is unlikely however, as the largest individuals from the S2 were all born on different days, in different tanks in the randomised design, and not treated any different than jars from the S1 or other treatments. Alternatively, it may be that the genetic basis of the response to selection is more variable in S-lines than in L-lines, causing greater variation in phenotypes between S-lines, an explanation which is clarified further in 2.5.3.

No divergence in standard length between selection treatments was observed in newborn fish. This may be explained by the measurements not having a sufficiently fine resolution to detect differences of such small magnitude. Alternatively, the results could reflect a true lack of response to selection of size at birth. This may indicate that differences in standard length of male guppies are a result of differences in growth rate, rather than physiological differences in body-plan or absolute growth. Furthermore, because no response to selection was observed in females, results indicate that the majority of the response is a result of Y-linked variation. As such variation does not necessarily get expressed in juveniles, a response to selection in juveniles may be absent. Alternatively, since juvenile measurements were done on both males and females, a response to selection in juvenile males may be concealed. However, as only adult standard length, and not growth rate was measured, these alternative explanations could not be tested here.

No response to selection was observed in female fish. On the contrary, although SL varied between generations as a result of variation in the mean age of females, mean SL was remarkably similar between the different selection lines, with the exception of a significant difference between the control and selection lines in the F5. This difference might indicate the process of selection itself, as opposed to random selecting the control, resulting in reduced female body size, possibly by increased inbreeding as a result of selection (van Oosterhout et al. 2007a). However, a difference between the control and treatments was observed in one generation only and is not confirmed by microsatellite data (see Chapter 3). More likely is, although levels of phenotypic variation were higher in females, and therefore small differences between lines would have been harder to detect, that such a lack of response to selection in females indicates strong Y-linkage of the genes underlying the response to selection observed in males. The findings by Tripathi et al. (2009a), who detected major QTLs for male size traits on the sex chromosome of their linkage map of guppies, support this.

Silliman (1975) also reported a response to selection in *Tilapia mossambica* males only: growth rate in males declined significantly over six generations of selective harvesting selection, but no response was evident in females. Furthermore, a delay in the response of females to changes in predation regime was also reported by Reznick and colleagues for their transplant experiments (Reznick and Bryga 1987; Reznick et al. 1997). Like the present study, these authors concluded that strong Y-linkage could explain the observed differences in response rate between the sexes, or alternatively that there might be actual differences in levels of genetic variance for this trait in males and females, or genetic covariance between body size and life history traits. Due to the nature of their experiment, they could not discriminate between the effects of selection on different (co-varying) aspects of life history and therefore could not explain or quantify the (lack of) response to selection on female body size directly. However, as only body size was under selection in the present study and environmental variation was minimal, the present study corroborates their suggestion of y-linkage of the genetic basis of body-size in male guppies.

### **2.5.2 Heritability of standard length**

Heritability estimates here were high; 0.518 overall and 0.538 and 0.398 for S-lines and L-lines independently and are comparable to estimates reported elsewhere. Inconveniently, in many studies addressing fisheries-induced evolution (e.g. Law 2000; Conover and Munch 2002; Law 2007), body length is treated as a life history trait whilst in meta-analyses of heritability in the wild, it is regarded as a morphological characteristic (Houle 1992; Weigensberg and Roff 1996; Kinnison and Hendry 2001). This complicates the transparent comparison of data. As body size affects fitness in many organisms and is closely related to size and age of maturation, fecundity and other such components of life history, it is understandable to regard body size itself as a life history characteristic. Strictly speaking however, it is a morphological trait.

The heritability estimates for standard length in this study are higher than those for life history traits as estimated in the meta-analysis of Weigensberg & Roff (1996), which were 0.268-0.323 on average. Estimated heritabilities for cod and salmonids

(including body length) were around the same magnitude ( $\pm 0.3$ ; Law 2000), and body length in Pacific Salmon was estimated at 0.4 (Funk et al. 2005). Heritability of morphological traits in Weigensberg & Roff (1996) was estimated at 0.501-0.562, which is similar to the estimates here. Reznick et al. (1997) estimated heritability of body size in male guppies (mg) at 0.88-0.998 which is much higher than our own estimates. Similar estimates ( $0.9 \pm 0.31$  s. e.) were found by Reynolds & Gross (1992) and compared to other estimates in guppies our estimates of heritability are therefore low. The reasons for the discrepancies in  $h^2$  estimates between current findings and these two prior studies may lie in population-specific differences in levels of y-linkage or otherwise be an effect of different rearing conditions in our respective experimental designs (D. N. Reznick, pers. comm.). Both Reznick et al. (1997) and Reynolds & Gross (1992) reared fish on food levels measured for each fish individually, while here we provided *ad libitum* amounts of food for whole tanks. Such free access to food may have induced differences in food intake between individual fish and consequently led to higher environmental variation.

However, the heritabilities estimated here are still likely to be exaggerated because our results and those elsewhere (Reynolds and Gross 1992; Reznick et al. 1997) indicate that male body size is, at least partially, Y-linked. Therefore, true heritability is probably lower since the correction for selecting on a single-sex is not necessary for fully Y-linked traits. For the same reason, Houde (1992) obtained extremely high estimates for heritability ( $h^2 > 1$ ) for orange colouration in guppies. Since we do not know the extent of Y-linkage of standard length in selection lines precisely, we cannot quantify such an effect, but it should be kept in mind that true  $h^2$  is likely to be lower than estimated here. Consequently, values of  $V_A$  may be overestimated. In addition, heritabilities from this experiment may be somewhat overestimated since no correction has been applied to account for the Bulmer effect (Bulmer 1971). This effect is the result of mass selection reducing variance between families, but not within families, caused by the “joined disequilibrium” of QTLs between related individuals (Bulmer 1971; Hill and Caballero 1992). Not accounting for this effect results in an upward bias of heritability estimated under strong selection. For the above reasons, the actual heritabilities observed here may be closer to those of life

history traits as estimated by e.g. Weigensberg and Roff (1996). However, as the effects of possible Y-linkage and the Bulmer effect have not been quantified here, this remains speculative.

### **2.5.3 Variance components and their response to selection**

Regardless of the possibility of heritability being overestimated in this study, notable differences in  $h^2$  were observed between up- and down selected lines. The explanation for this might be found in several different directions.

Heritability represents a link between additive genetic and phenotypic variation. It cannot be excluded that discrepancies in the accuracy of measurements and the selection procedure resulted in mismatches between observed and expected levels of variance reduction. Additionally, the consistent slower than expected rate of decline in  $V_p$  in selection lines in both directions indicates some source of  $V_p$  is unaccounted for in our results. However, this cannot explain the strong difference in observed levels of phenotypic variation between S- and L-lines. Therefore, reduced heritability in L-lines indicates either lower  $V_A$  or higher levels of  $V_R$  (Falconer and MacKay 1996) compared to S-lines. As  $V_A$  was similar to- or even slightly higher than expected and no decrease in  $V_p$  was observed in L-lines, the lower  $h^2$  in these lines is most likely the result of elevated levels of  $V_R$ . Although all fish shared the same environment and were reared under identical conditions, some source of environmental variation may differ between treatments that may explain the higher environmental variation for L-lines than for S-lines. For example, slight variations in density or proportionally higher intake of food of the largest individuals in L-lines may have accentuated differences in food intake between fish in these lines. Therefore, although this is a speculative argument, this elevated environmental variation may account for some of the difference in  $V_p$  levels between S- and L-lines.

Alternatively, levels of non-additive genetic variation may be different for the differently selected lines, i.e. dominance and interaction variance. In particular the combined observation of high  $V_p$  and higher than expected  $V_A$  in L-lines, compared to the resemblances in observed and expected levels of  $V_A$  in S-lines, may reflect

changes in levels of dominance variance in L-lines. Dominance variation can be expected to differ between up- and down selected fish, and this can be explained by the possible differences between the genetic bases of large and small body size.

As a quantitative trait, the genetic basis of body size is the product of many genes and therefore non-additive genetic variation can be expected to contribute to trait variation through dominance and epistatic effects. It is not possible here to make any inferences on the effect of genetic interactions between loci on our results. However, large body size is often more closely related to fitness, for example due to improved body condition (Brown et al. 1993) or female choice (Barbosa and Magurran 2006). Although evidence for female preference for large males is inconclusive in guppies (e.g. Reynolds and Gross 1992; Endler and Houde 1995; Evans et al. 2003; Magellan et al. 2005) the observation of increased male size in the absence of predation at least suggest that guppies are no exception to this rule and large male body size may be closer related to fitness than small size, as field experiments show (Reznick et al. 1997).

In his meta-analysis of evolvabilities of traits, Houle (1992) observed higher levels of residual variance for traits closely related to fitness. Such traits, he reasoned, are the combined result of all the selective forces acting on their lower-level components. Consequentially, a larger number of interactions and variables will affect such fitness-related traits. However, being closer related to fitness, large body size is also more likely to be the specific end product of a confined number of pathways, whilst small body size can be achieved through more diverse ways, since reduced size is a more general effect of recessive mutations (Wright 1977; Charlesworth and Charlesworth 1999). Reduced body size as a result of inbreeding for example, has been reported for a range of species (see e.g. Falconer and MacKay 1996 for references) including guppies (van Oosterhout et al. 2007a) and in lake whitefish species (*Coregonus* spp.), markedly lower levels of heterozygosity were observed for dwarf phenotypes than normal-sized fish (Renaut et al. 2011). Such a broader genetic basis to small size may also explain the greater variation in the selection response of maturation between S-lines compared to L-lines.



At least two different processes may therefore contribute to explaining the increased  $V_R$  in large, but not in small fish in the present study. First of all, the closer association with fitness of large body size may contribute to elevated levels of  $V_R$  in large fish. Secondly, small body size can be expected to be associated with higher levels of recessive homozygotes than large size, and consequently reduced dominance effects. Following the same rationale, if large body size is associated with elevated levels of heterozygosity, this should maximise  $V_D$  and contribute to the high variance observed in L-lines. Both rationales, as well potential variation in  $V_R$  between lines due to birth rates and food competition, can explain our data and the observation of elevated levels of non-additive variation in L-lines compared to S-lines.

#### **2.5.4 Evolutionary rate of standard length**

##### **Natural rates of microevolution**

How do the observed rates of evolvability compare to those observed elsewhere? Compared to estimates from wild populations, the estimates of evolutionary rate from our experiments ( $50\text{-}55 \times 10^3$  d or 0.2-0.3 haldanes) are very fast. Kinnison and Hendry (2001) observed median rates of 0.035 haldanes in their meta-analysis of studies on contemporary evolution spanning less than 80 generations. Over their entire analysis (over 2000 estimates from studies spanning up to 140 generations) average evolutionary rates were 0.0058 haldanes and  $1.2 \times 10^3$  d; our rates are ten times higher when calculated in haldanes for the short-generation time studies and around 45 (darwins) and 35-50 (haldanes) times higher when compared to their entire study. Of a greater magnitude than those from Kinnison and Hendry's meta-analysis, were the rates observed after transplant experiments with wild guppies: male size (mg) changed at a rate of  $5.3\text{-}27.1 \times 10^3$  d over 7-18 generations (Reznick et al. 1997) or 0.106 haldanes (Reznick and Ghalambor 2005a): 0.2-0.5 times the rate we observed. Darimont et al. (2009) estimated different evolutionary rates for anthropogenically perturbed and naturally evolving systems, and found that the

former increased evolutionary rates by 50-300%. Rates in  $10^3$  d averaged at 12.06 ( $\pm$  10.45 st.dev) and were highest for marine harvested fish species.

### **Experimental rates of microevolution**

When comparing our results to other instances of artificial selection, the results are very similar. Gingerich (1983) reported an average evolutionary rate of  $58.7 \times 10^3$  d (range 12-200  $\times 10^3$ ) from artificial selection experiments, which is close to what was observed in the present study. Using the figures in their publication, we estimated the rate of evolutionary change in darwins for the selection experiments by Conover and Munch (2002). Mean weight of harvested fish (g) increased at a rate in the order of  $156 \times 10^3$ , and decreased at approximately  $95 \times 10^3$  d (using an estimated generation time of 235 days). As their selection intensity was much higher (10% of individuals and on both sexes) these double-to-triple rates of evolution when compared to our own work are in the line of what one would expect.

### **Reasons for high observed rates of evolution**

Our high evolutionary rates can in the first place be explained by the high selection intensity and non-overlapping generations that are intrinsic to the experimental design. Furthermore, generation time in the present study was short (6 months) and rates of evolution observed tend to be higher for studies with short generation time (Kinnison and Hendry 2001) or those estimated over brief time periods (Gingerich 1983). Another possibility is that the constant environment in our experiments resulted in high  $h^2$  and consequently rapid evolution of traits. However, although theory suggests  $h^2$  to be affected by environmental conditions (Falconer and Mackay 1996; Hoffmann and Merilä 1999), a considerable body of empirical data suggests  $h^2$  does not differ between laboratory and natural conditions (Weigensberg and Roff 1996; Stokes and Law 2000; Conover and Munch 2002). The even higher heritability for body size in guppies observed by Reynolds & Gross (1992) and Reznick et al. (1997) suggest furthermore that our estimates are not extreme. I am therefore inclined to regard the differences in evolutionary rate observed between our work and natural systems as an effect of the selection intensity employed, rather than of elevated levels of  $V_A$  of the trait under selection in our work compared to others.

The evolutionary rates observed here thus are similar to those reported for other artificial selection systems, but much higher than those observed in natural systems. However, our estimates are less than five times as high than those for human-harvested systems estimated by Darimont et al. (2009) which is, considering the strong selection regime imposed here, painfully illustrative of the potential impact human harvesting can have on natural systems.

### **2.5.5 Overall conclusions**

The work reported here has demonstrated that size-selective harvesting is capable of inducing rapid shifts in correlated morphological and life history traits in guppies and provides further advances in our understanding of life history evolution in general. It was known already from field experiments (Endler 1980; Reznick and Bryga 1987; Reznick et al. 1990) that guppies are capable of rapid adaptation to a release of size-selective predation. The present experiments confirm this observation. Furthermore, as a high predation population was used to establish our selection lines, and based on the symmetrical nature of the response, the present study demonstrates that a history of high predation and adaptation hereto in the wild has not deprived guppies of additive genetic variation for life history traits. Although only a single wild population was used, and therefore care must be taken with inferring general conclusions from our observations, we have demonstrated a considerable additive genetic component to variation in standard length and confirmed that the rapid response to selection is at least partially due to Y-linkage of size-related traits in male guppies. Furthermore, our data suggests that large body size in guppies is partially governed by dominance variation at the underlying loci, whilst small fish possess less non-additive genetic variation and are more likely to be the result of greater occurrence of recessive homozygotes.

To what extent can the present study inform on fisheries-induced evolution? Previous experimental work (Conover & Munch 2002) with the same aim has been scrutinised for having unrealistically strong selection intensities. In addition, non-overlapping generations, the knife-edged aspect of the experimental design and the absence of compensatory processes made the observed response not directly

comparable to commercial fisheries (Hilborn and Minte-Vera 2008). Similar criticisms can be justly delivered here: our system is simplified and does not incorporate realistic fishing pressures or compensatory processes. However, we maintain that our findings do nevertheless make a valuable contribution to the ongoing debate on the importance of fisheries-induced evolution for various reasons.

Firstly, the knife-edge selection regime was required here to make it feasible to obtain a response within the duration of this study. Whilst unrealistic in its current form, it is possible to transform the data obtained here to more realistic time spans and selective pressures (Brown et al. 2008).

Secondly, some of the most convincing examples on fisheries-induced evolution come from semelparous Pacific salmon (Ricker 1981) and there is substantial concern over the evolutionary effects of fishing on salmon stocks, with fishing mortalities as high as 70% (Ricker et al. 1978) and 80-90% (Hard et al. 2008) reported. The evolutionary rates observed here may therefore be directly comparable to rates observed in salmon and provide support for the occurrence of evolutionary change in exploited salmon stocks. In addition, the symmetry of the response we observed from a population that had a history of intense size-selective predation in the wild bodes well for the reversibility of an evolutionary response to selection, though empirical testing is required here.

Thirdly, whilst the absence of realistic density- and compensatory processes in our design is likely to have affected the results, this simplified design maximises the likelihood of picking up a genetic response to selection and sheds light on the genetic basis of body size and the dynamics between selection intensity and its genetic and phenotypic response. In light of fisheries-induced evolution, the principles shown here demonstrate the directionality and potential for selection on body size to induce detectable genetic change over few generations of selection. We have clearly demonstrated that there is an important interplay between the selection intensity, phenotypic and genetic change and it is highly likely that this is at least partially illustrative of the dynamics of fish and fisheries on a greater scale.

3.

GENETIC DYNAMICS OF

SELECTION LINES OF *P. RETICULATA* I:

GENETIC VARIATION AT NEUTRAL

MICROSATELLITE LOCI AND DRIFT BETWEEN

SELECTION LINES

### **3.1 INTRODUCTION**

#### **3.1.1 The use of microsatellites for studying genetic diversity of exploited fish species**

Microsatellite markers have been used widely to answer an array of questions on the dynamics and connectivity of wild populations (Awise 2004). For aquatic species in particular, whose migratory behaviour and intraspecific interactions occur under water, and whose abundance can easily conceal complex structure (Conover et al. 2006), molecular markers are of great use. For example, neutral microsatellite markers have been previously used in guppies (see 1.4.2) and a wide range of marine and freshwater fish species to assess levels of genetic variation (DeWoody and Awise 2000), expose spawning competition (Bekkevold et al. 2002) and correlations between offspring and parental traits through paternity analysis (Johnson et al. 2011), reveal population structure (Ruzzante et al. 2000; Bernal-Ramírez et al. 2003) and migration patterns (Hendry 2000; Ruzzante et al. 2006). In addition, microsatellites have been very valuable for population assignment of mixed stocks (Nielsen et al. 2001) and estimating effective population sizes (Hauser et al. 2002; Poulsen et al. 2006), which, molecular markers revealed, can be orders of magnitude lower than census sizes. More recently, dense mapping of microsatellites and the design of expressed sequence tag (EST)-linked microsatellites has provided evidence of putative non-neutrality and selective responses of microsatellites (Nielsen et al. 2006; Larsson et al. 2007), making them useful markers in adaptation studies using genome scans and outlier analyses (Hansen et al. 2010; Martínez et al. 2011).

Microsatellite markers have also proven useful for studying evolutionary effects of fishing (Hauser and Seeb 2008), although this field has, with few exceptions (Carvalho and Pitcher 1994; Park and Moran 1994), only emerged recently. The majority of genetic work in fisheries, besides advances in the genetics of aquaculture breeding (De-Santis and Jerry 2007), has focused on estimating effects of exploitation on neutral genetic diversity, and reductions in effective population size ( $N_e$ ) have been associated with harvesting pressure (Hauser et al. 2002; Hutchinson et al. 2003; Hoarau et al. 2005).

Such reductions in neutral genetic diversity associated with fishing pressure, are virtually the only genetic data currently available in support of fisheries-induced evolution. However, in most exploited species, even very low  $N_e:N$  ratios still imply thousand(s) or more breeding individuals, and such skewed ratios are observed in combination with temporal stability of allele frequencies (Poulsen et al. 2006; Cuveliers et al. 2011). Loss of neutral genetic diversity through drift therefore may not be an immediate concern for many intensely harvested stocks (Poulsen et al. 2006, Cuveliers et al. 2011). In large marine fish populations, drift typically operates at much lower rates than selection (Wright 1951). Therefore, neutral genetic variation may be especially unsuitable as a measure of evolutionary change as specific traits affected by fisheries selection, such as growth and aspects of life history, are likely to evolve at rates that differ from those of neutral genetic markers (Merilä and Crnokrak 2001; McKay and Latta 2002; Luttikhuizen et al. 2003; Hutchings et al. 2007). In light of fisheries-induced evolution (FIE), neutral genetic markers may therefore be best suited as a genetic baseline to compare putative candidate markers to. This was, for example, done by Jakobsdóttir et al. (2011), who observed temporal stability at six microsatellite loci, compared to strong changes in allele frequencies at the *Pan I* locus that could be associated with both fishing pressure and changes in age structure in cod. Comparing such baseline variation to candidate gene frequencies offers an opportunity for assessing the effects of selection at a molecular level (Hemmer-Hansen et al. 2007; Hansen et al. 2010; Tonteri et al. 2010). Moreover, monitoring of neutral genetic diversity over time or across generations can reveal information on mating patterns and inbreeding.

### **3.1.2 The use of microsatellites for studying genetic diversity in artificial selection experiments**

In contrast to many natural fish populations, drift may be of considerable importance in small confined populations, e.g. captive populations and selection lines, where furthermore inbreeding and mating patterns may greatly affect levels of genetic variation. For this reason, neutral genetic variation needs to be monitored, quantified and valued correctly.

Surprisingly however, with the exception of studies focusing on bottleneck effects on genetic diversity (Montgomery et al. 2000; England et al. 2003) only few studies have specifically monitored neutral genetic variation in artificial selection experiments (Morgan et al. 2003; Barrett et al. 2010) and even fewer have done so in combination with exploring putative candidate gene frequencies. Numerous experimental studies use molecular markers for e.g. QTL-mapping or genome scans to identify candidate loci of interest (Sørensen et al. 2007; Rand et al. 2010; Svetec et al. 2011), but I found none that actually compared the dynamics of neutral and putative selected loci directly. By far the majority of studies focuses solely on candidate loci and fails to report comparisons with neutral loci (Tao and Boulding 2003; Belter et al. 2004; Fidler et al. 2007). However, high levels of drift and/ or loss of genetic variability in selection lines can be a concern since high levels of genetic erosion and drift could lead to a loss in adaptive variation and/ or conceal the effects of selection (Hartl and Clark 1989). Several studies have concluded that selective breeding, even in small groups (<15 breeding pairs per generation), did not lead to a loss in genetic variability at neutral microsatellites loci over few generations, but did however result in significant drift of allele frequencies between selection lines and within the same line over generations (Morgan et al. 2003; Simões et al. 2008a, 2008b, 2010). Considering the extent of work that has been building on selection experiments (Chippindale 2006), endeavours towards a thorough understanding of molecular genetic dynamics of both neutral markers and putative candidates for selection in experiments, seem justified.

For the present study, microsatellites were used to contribute to two different aspects of genetic assessment: (1) to provide a baseline to compare putative candidate gene frequencies to and (2) to quantify the effects of drift and artificial selection in small, captive breeding populations and selection lines on supposedly neutral marker loci.



## **3.2 MATERIAL AND METHODS**

### **3.2.1. Origin of fish and rearing protocol**

All males used for breeding in the F0-F5 generations (50 males per line and generation) and those used for estimating maturation size and age in the F6 generation were used for genetic analysis. Rearing- and selection protocols can be found in section 2.3.

### **3.2.2. DNA extraction**

Genomic DNA was extracted using hexadecyltrimethylammonium bromide (CTAB) according to the following protocol: 2-3 mm<sup>2</sup> of tissue was incubated overnight at 60°C in 350µl 2% CTAB buffer (100mM Tris-HCl pH 8.0; 1.4M NaCl; 20mM EDTA pH 8.0; 2% CTAB; 0.2% 2-mercaptoethanol) with 20µl proteinase K solution (QIAGEN). For the extraction, 300µl chloroform: isoamyl alcohol (24:1) was added, mixed for 5 min using an automated rotator and then centrifuged at 13.000rpm for 5min. The supernatant was transferred to a new 1.5ml tube and the extraction step repeated, after which 660µl 100% ethanol and 30µl 3M NaAc pH4.8 were added to the retained supernatant. This solution was mixed for 3 min, left to stand for 10 min and then centrifuged at 13.000rpm for 10 min, after which the supernatant was discarded and 500µl 70% ethanol was added to the pellet. This was centrifuged for 5 min to wash the pellet and the supernatant was discarded again. Pellets were left to dry in a 37°C cabinet and re-suspended in 100µl H<sub>2</sub>O.

DNA concentrations were quantified using a NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies) and ranged from 6-157ng/µl. A working solution of DNA was stored at 4°C and diluted to a volume of 50µl at 10ng/µl for samples with initial DNA concentrations over 10ng/µl. Remaining DNA was kept at extraction concentration and stored at -20°C.

### **3.2.3 Microsatellite marker choice and PCR protocols**

To monitor neutral genetic dynamics in selection lines over generations, male fish were genotyped at 10 microsatellite loci: Pr39, Pr92 (Becher et al. 2002), Pret-32,

Pret-69, Pret-77 (Watanabe et al., 2003), AGAT11 (Olendorf et al. 2004), Hull70-2 (van Oosterhout et al. 2006a), G82, G102 and G289 (Shen et al. 2007a). Markers were chosen such that the entire marker set comprised markers with different numbers of alleles and minimal overlap of allelic ranges with other chosen loci. Forward primers were extended with an 8bp tail, complementary to a FAM/NED/VIC/PET fluorescent dye-labelled universal primer following Schuelke (2000; Tysklind 2009) and amplified using QIAGEN Multiplex PCR kit. All primer sequences are given in Appendix I.

Annealing temperatures (AT) were 53° for Pret-69, Pret-77, Pr39, G102, Hull70-2 and AGAT11 and 58° for Pret-32, G82, Pr92 and G289 and primers with the same AT were amplified in a single 10µl reaction: 6µl Multiplex mix, 1µl Q solution, 1µl H<sub>2</sub>O, 1µl DNA at 10ng/µl and 1µl primer mix (containing a mix of equal volumes of forward primers at 1µM and reverse primers and dye-labelled universal primers at 10µM concentrations). PCR products were obtained using a Tetrad2 Peltier thermal cycler (BIO-RAD) and the following program: 95°C 15m; 94°C 30s, AT 90s, 72°C 90s for 12 cycles; 94°C 30s, 50°C 90s, 72°C 90s for 30cycles; 60°C 30m.

PCR products were resolved on a 3130xl Genetic Analyzer (Applied Biosystems) using GeneScan™ LIZ® 500 as an internal size standard and genotyped using GeneMapper® Software v.4.0 (Applied Biosystems). Genotyping accuracy was verified by re-analysing several individuals in multiple runs and by running negative controls on each plate.

### **3.2.4 Genetic analyses**

#### **Terminology**

All genetic analyses were performed for each generation and for each selection line independently. In the F3 generation, all sampled individuals were part of the same breeding group, i.e. they were reared under identical conditions and shared the

same parents. However, as the samples were divided over lines in a non-random fashion (one randomly selected line and four lines consisting of the extreme 20% percentiles of the generation, see 2.3.5), the five lines of the F3 generation were treated as separate breeding groups like all selection lines from the F4-F6 generations throughout the analyses. Thus, in total, 23 different breeding groups were analysed (F0 – F1 – F2 – 5 x F3 – 5 x F4 – 5 x F5 – 5 x F6) which from now on will be referred to as either lines (over generations) or populations (within generations).

### **Null alleles, Hardy-Weinberg equilibrium and linkage**

All microsatellites were checked for null alleles using MICRO-CHECKER version 2.2.3 (van Oosterhout et al. 2004) and confirmed by  $G_{IS}$   $p$ -values, obtained using 9999 permutations. Benjamini & Hochberg (B-H) corrections for multiple tests were performed using SGoF+ (Carvajal-Rodríguez et al. 2009) here and throughout the analyses.

Conformity to Hardy-Weinberg (HW) equilibrium was tested at each locus in each population and for each generation independently, using the Hardy-Weinberg probability test in Genepop version 4.0.10 (Raymond and Rousset 1995; Rousset 2008) using 10.000 batches and 10.000 iterations per batch. Results were corrected for multiple tests as above.

Genotypic linkage disequilibrium (LD) was tested for using Genepop's log likelihood ratio statistic for each pair of loci in each population. LD was only tested for in the wild-caught (F0) fish because (artificial) selection increases the relatedness between individuals, and thereby affects the independent inheritance of alleles between individuals. This affects all alleles, and is not restricted to alleles at loci that are physically linked (Falconer and Mackay 1996). Hence, artificial selection is predicted to increase LD irrespective of the genomic distance between marker loci. In other words, physically linked loci are expected to remain in LD throughout the experiment, whereas alleles of unlinked loci may become in LD without any physical (genomic) linkage in later generations.

### **Genetic diversity**

Using GenoDive (Meirmans and van Tienderen 2004), the following parameters were obtained for each selection line and for each generation independently, per locus and over all loci; number of alleles  $A$ , observed heterozygosity  $H_o$ , expected heterozygosity  $H_s$ , and inbreeding coefficient  $G_{IS}$  (Nei 1987). Standard errors over all loci were obtained through jackknifing over loci and  $p$ -Values for  $G_{IS}$  were obtained using 9999 permutations.

Allelic Richness  $A_R$  (El Mousadik and Petit 1996) was calculated per locus based on a minimum sample size of  $N=32$  (F1), using FSTAT version 2.9.3.2 (Goudet 1995). Significant differences in  $A_R$  within generations were tested for using nested ANOVA (PASW Statistics 18.0.3). To test for significant changes in  $A_R$  over generations, randomisation tests with 10.000 randomisations were performed using Rndom Pro 3.14 (Jadwiszczak 2009).

### **Genetic Structure**

To examine whether variation between populations was more likely to be a result of selection or random drift between lines, hierarchical analyses of molecular variance (AMOVA) were performed for each generation independently, using Arlequin version 3.5.1.2 (Excoffier et al. 2005). AMOVAs were carried out for each generation (F3-F6) independently in a hierarchical fashion using the following levels: (1) within lines, (2) within treatments-between replicate lines, and (3) between treatments. These hierarchical levels are equivalent to  $F_{ST}$ ,  $F_{SC}$  and  $F_{CT}$  respectively. In addition, an AMOVA was carried using the following hierarchical levels: within lines within generations ( $F_{ST}$ ), within lines between generations ( $F_{SC}$ ) and between lines ( $F_{CT}$ ). This hierarchy was used to examine whether temporal differentiation within lines or differentiation between lines explained most of the variation.

In addition, both pairwise  $F_{ST}$  estimator  $\theta$  (Weir and Cockerham 1984) and standardised fixation index  $G''_{ST}$  (Meirmans and Hedrick 2011) were calculated per generation for each locus for within-treatment comparisons (S1-S2 and L1-L2),

between-treatment comparisons (S1-L1, S1-L2, S2-L1 and S2-L2) and comparisons with the control line. (C-S1, C-S2, C-L1 and C-L2) using GenoDive. *P*-Values were obtained for pairwise  $F_{ST}$  comparisons over all loci only, using 9999 permutations, for testing significance between treatments by nested ANOVA and over generations by randomisation tests (10.000 randomisations).

### **3.2.5 Paternity analysis**

To estimate the number of fathers and male reproductive variance in each population, paternity of males was reconstructed using CERVUS version 3.0.3 (Marshall et al. 1998; Kalinowski et al. 2007). Only father-son pairs that were assigned with strict (95%) confidence and no mismatching loci were considered and used for estimating male reproductive success and father-offspring regression of standard length. The slope of the linear relationship between the deviation of offspring standard length from population means on father standard length from population means, is furthermore an estimate of heritability (Falconer and Mackay 1996).

### **3.3 RESULTS**

#### **3.3.1 Loci amplification**

All ten microsatellite loci amplified in all selection lines. For locus AGAT11, a variable number of individuals was observed with three clear peaks of almost identical height in seven out of the 23 populations tested. Re-amplification of PCR products and re-genotyping yielded the same results. The third peak was not consistently the same allele in different individuals, nor linked to one other specific allele, although most of the individuals with three peaks carried an allele at 332bp. Re-analysing the data set without allele 332 resulted in high deficiencies in heterozygosity in most of the samples. Paternity analysis indicated that the third allele was inherited (most three-peak individuals had a father with three alleles as well) but a straightforward inheritance pattern could not be identified. The frequency of individuals with three alleles in lines with the allele present varied from 6-14%, but an individual with more than three peaks was never observed. In all likelihood, AGAT11 is located on the (pseudo) y-chromosome and some males may carry an additional copy of this locus. Although no genetic samples for females are available, and sex-linked inheritance could therefore not be tested, this would explain the lack of individuals with four alleles for this locus. Although none of the populations with three-peak individuals had significant deviations from HW-equilibrium, the locus was excluded from further analyses.

#### **3.3.2 Null alleles, Hardy-Weinberg equilibrium and linkage**

##### **Null alleles**

For three out of nine loci, MICRO-CHECKER indicated that null alleles could explain observed homozygote excess in one out of 23 tests per locus (Hull70-2 in the F2, G102 in F3\_L2 and Pret-77 in F6\_C). For locus Pret-32, null alleles were indicated as explanations for HW-deviations in the F1 and F4\_S2 populations. For one locus, G82, null alleles were highlighted as explaining homozygote excess in three tests (F2, F5\_L1 and F6\_C). However, the homozygote excess in these populations was not

supported by significant values of  $G_{IS}$  after B-H correction (see  $G_{IS}$  section below). Furthermore, all 23 populations were derived from the same gene pool very recently and null alleles present in one population should therefore be expected to be present in a majority of the populations. Finally, one of the assumptions of MICRO-CHECKER was violated from the F3 onwards, since artificial selection was done, which could result in deviation from HW-equilibrium. Therefore, the observed homozygote excess is more likely the result of stochastic variation in allele frequencies over generations and an inflated type I error rate due to multiple tests, rather than an indication that null alleles were present at these loci.

### **Hardy-Weinberg equilibrium**

Tests for HW-equilibrium indicated deviations in 8 out of 30 tests in the random breeding generations (F0-F2) and 22 out of 200 tests in selection lines (F3-F6). After Benjamini-Hochberg correction, a highly significant ( $p < 0.001$ ), heterozygote excess remained only for locus Pr39 in F5\_L2, F6\_L1 and F6\_L2 and this locus was removed from further analysis with neutral microsatellite loci. However, due to the consistency of this deviation with the selection regime (L-lines only), which may indicate selection is acting at this locus or a closely linked locus, locus Pr39 was further analysed with the putative candidate loci and is discussed in Chapter 4.

### **Linkage Disequilibrium**

Four out of 36 tests demonstrated significant ( $p < 0.05$ ) linkage disequilibrium between loci (Pret-32 with locus G102 and G82 and locus Pr92 with locus G289 and G82). However, none of these were significant after B-H corrections for multiple tests.

### **$G_{IS}$**

$G_{IS}$  was calculated per locus and over all loci (Fig. 3.1) for each population independently. Per locus, 34 out of 207 tests yielded significant deviations of  $G_{IS}$  from zero. After B-H corrections, only locus Pr39 in the F6\_L1 remained significant ( $G_{IS} = -0.381$ ).  $G_{IS}$  values therefore do not support the conclusion of null alleles by homozygote excess from MICRO-CHECKER, lending further support to my

interpretation that the type I error rate was inflated due to performing multiple tests with MICRO-CHECKER. Over all loci, 6 out of 23 (including Pr39) and 4 out of 23 (excluding Pr39) tests had significant values of  $G_{IS}$ , none of which were significant after B-H corrections.

### **3.3.3 Genetic diversity**

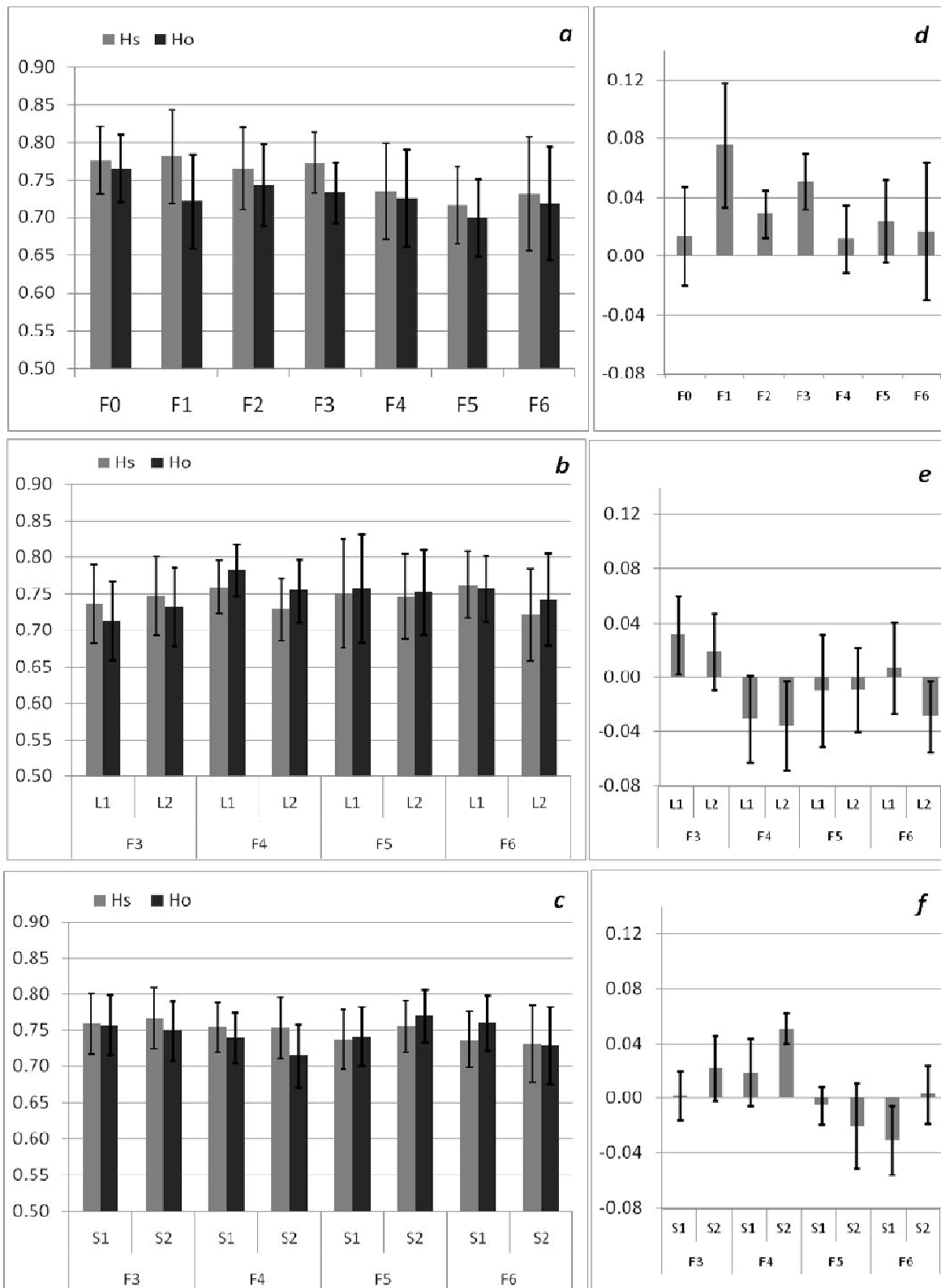
#### **Heterozygosity**

Genetic diversity within samples was moderate to high with heterozygosity levels ranging from 0.7-0.78. Only slight variations in heterozygosity estimators were observed between populations, both between treatments and lines, and over generations (Fig. 3.1), and these were not significant ( $p > 0.171$  and  $p > 0.09$  for all within-generation nested ANOVA comparisons of  $H_O$  and  $H_S$ , and  $p > 0.243$  for all regressions over generations, either over all lines or within lines). In generations F5 and F6, a general excess of heterozygotes was observed in both S- and L-lines, but not in the control (Fig. 3.1), but this excess was not significant (see  $G_{IS}$  above). A summary of population genetic parameters is given in Table 3.1. A table of allele frequencies at microsatellite loci can be found in Appendix II.

#### **Allelic Richness**

The numbers of alleles ranged from 3-29 depending on the locus (Table 3.2) and did not differ between selection lines (nested ANOVA over all loci per generation,  $p > 0.789$  for each generation). Allelic Richness ( $A_R$ , Fig. 3.2) was significantly higher in the F0-F1 generations (10.5-10.35) than in the captive-bred generations F2-F6 (mean  $A_r = 8.44$ , randomisation test,  $p = 0.032$ ). No significant decline in  $A_R$  was observed in generations F2-F6 (logarithmic regression,  $R^2 = 0.026$ ,  $F_{1,134} = 3.593$ ,  $p = 0.06$ ) over all treatments, or in the random breeding control line ( $R^2 = 0.054$ ,  $F_{1,38} = 2.151$ ,  $p = 0.15$ ). Furthermore, for selection lines in generations F3-F6, no decline in  $A_R$  was observed ( $p = 0.52, 0.57, 0.58$  and  $0.45$  for the S1, S2, L1 and L2 respectively).

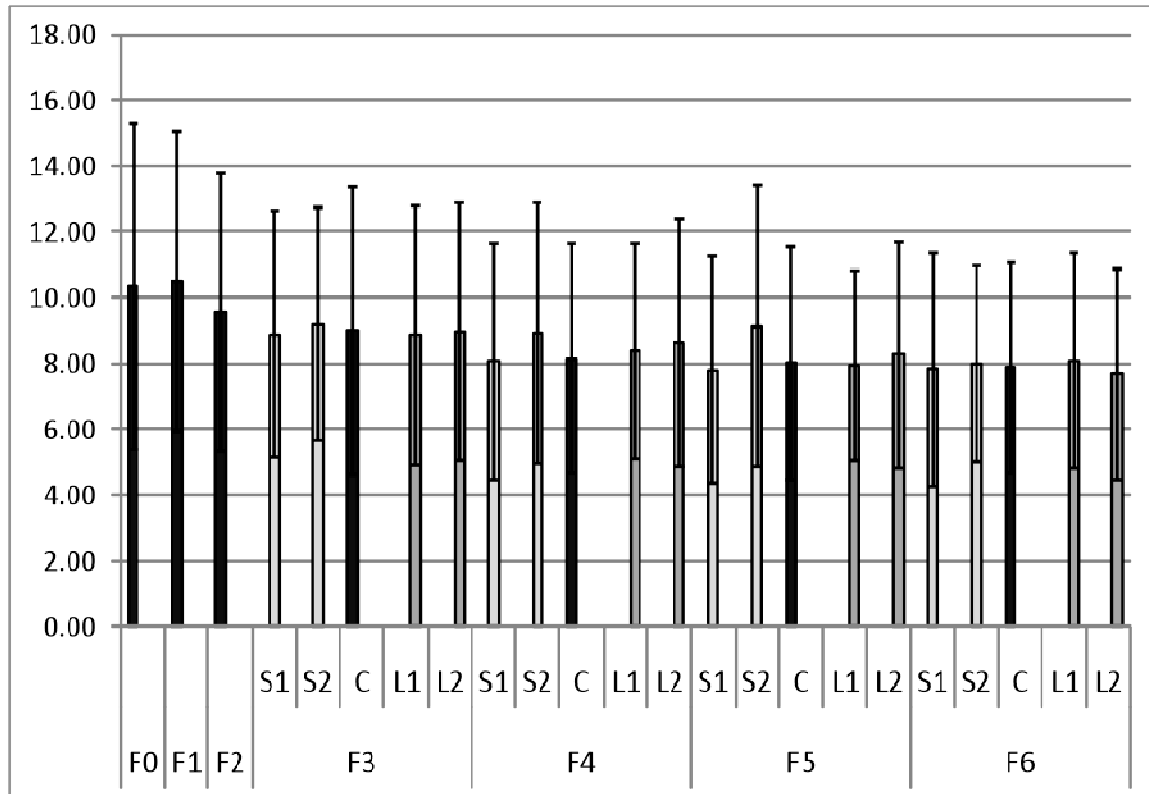




**Fig. 3.1.** Expected ( $H_s$ , grey bars) and observed ( $H_o$ , black bars) heterozygosity per generation over eight microsatellite loci for the random breeding control line (a), L-lines (b) and S-lines (c). Frames d-f provide  $G_S$  estimates for C-, L- and S-lines respectively. Error bars indicate standard errors, obtained through jackknifing over loci. None of the  $G_S$  values were significantly higher than zero after Benjamini-Hochberg correction.

**Table. 3.1.** Overview of population genetic parameters observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_s$ ) and inbreeding coefficient  $G_{IS}$ , number of alleles per locus ( $A$ ) and Allelic Richness ( $A_R$ ) over eight microsatellite loci. Standard errors were obtained through jackknifing over loci. S-lines are indicated by light grey shading, L-lines by dark grey, and random breeding populations are white.

<b>Population</b>	<b>N</b>	<b><math>H_o \pm St. err</math></b>	<b><math>H_s \pm St. err.</math></b>	<b><math>G_{IS} \pm St. err.</math></b>	<b><math>A \pm St. dev.</math></b>	<b><math>A_R \pm St. dev.</math></b>
<b>F0</b>	46	0.7658 $\pm$ 0.044	0.7762 $\pm$ 0.044	0.014 $\pm$ 0.039	11.50 $\pm$ 5.63	10.35 $\pm$ 4.92
<b>F1</b>	32	0.7227 $\pm$ 0.062	0.7814 $\pm$ 0.042	0.076 $\pm$ 0.042	10.50 $\pm$ 4.54	10.50 $\pm$ 4.54
<b>F2</b>	192	0.7435 $\pm$ 0.054	0.7657 $\pm$ 0.047	0.029 $\pm$ 0.016	11.88 $\pm$ 5.44	9.56 $\pm$ 4.19
<b>F3_S1</b>	52	0.7572 $\pm$ 0.042	0.7589 $\pm$ 0.047	0.002 $\pm$ 0.018	9.63 $\pm$ 4.10	8.87 $\pm$ 3.74
<b>F3_S2</b>	45	0.7500 $\pm$ 0.042	0.7668 $\pm$ 0.042	0.022 $\pm$ 0.024	9.88 $\pm$ 3.91	9.19 $\pm$ 3.56
<b>F3_C</b>	47	0.7336 $\pm$ 0.040	0.7729 $\pm$ 0.045	0.051 $\pm$ 0.019	9.50 $\pm$ 5.04	9.01 $\pm$ 4.38
<b>F3_L1</b>	49	0.7143 $\pm$ 0.054	0.7366 $\pm$ 0.058	0.031 $\pm$ 0.029	9.63 $\pm$ 4.34	8.87 $\pm$ 3.94
<b>F3_L2</b>	50	0.7325 $\pm$ 0.054	0.7467 $\pm$ 0.052	0.019 $\pm$ 0.028	9.75 $\pm$ 4.40	8.96 $\pm$ 3.91
<b>F4_S1</b>	50	0.7400 $\pm$ 0.035	0.7543 $\pm$ 0.037	0.019 $\pm$ 0.025	8.75 $\pm$ 4.17	8.08 $\pm$ 3.59
<b>F4_S2</b>	50	0.7150 $\pm$ 0.043	0.7533 $\pm$ 0.049	0.051 $\pm$ 0.011	9.75 $\pm$ 4.33	8.90 $\pm$ 3.97
<b>F4_C</b>	48	0.7266 $\pm$ 0.064	0.7350 $\pm$ 0.054	0.012 $\pm$ 0.023	8.75 $\pm$ 3.88	8.15 $\pm$ 3.50
<b>F4_L1</b>	49	0.7830 $\pm$ 0.036	0.7598 $\pm$ 0.046	-0.031 $\pm$ 0.032	8.88 $\pm$ 3.44	8.38 $\pm$ 3.26
<b>F4_L2</b>	49	0.7551 $\pm$ 0.043	0.7290 $\pm$ 0.053	-0.036 $\pm$ 0.033	9.38 $\pm$ 4.21	8.64 $\pm$ 3.76
<b>F5_S1</b>	46	0.7418 $\pm$ 0.041	0.7380 $\pm$ 0.042	-0.005 $\pm$ 0.014	8.38 $\pm$ 3.78	7.80 $\pm$ 3.46
<b>F5_S2</b>	43	0.7703 $\pm$ 0.036	0.7554 $\pm$ 0.051	-0.020 $\pm$ 0.031	9.75 $\pm$ 4.80	9.12 $\pm$ 4.28
<b>F5_C</b>	43	0.6998 $\pm$ 0.051	0.7171 $\pm$ 0.054	0.024 $\pm$ 0.028	8.38 $\pm$ 3.85	8.01 $\pm$ 3.53
<b>F5_L1</b>	46	0.7582 $\pm$ 0.075	0.7510 $\pm$ 0.050	-0.010 $\pm$ 0.041	8.38 $\pm$ 3.25	7.94 $\pm$ 2.88
<b>F5_L2</b>	49	0.7526 $\pm$ 0.058	0.7462 $\pm$ 0.047	-0.009 $\pm$ 0.031	8.88 $\pm$ 3.83	8.27 $\pm$ 3.45
<b>F6_S1</b>	50	0.7600 $\pm$ 0.039	0.7373 $\pm$ 0.043	-0.031 $\pm$ 0.025	8.38 $\pm$ 3.89	7.84 $\pm$ 3.55
<b>F6_S2</b>	50	0.7300 $\pm$ 0.0533	0.7321 $\pm$ 0.051	0.003 $\pm$ 0.021	8.63 $\pm$ 3.25	7.97 $\pm$ 2.98
<b>F6_C</b>	50	0.7196 $\pm$ 0.075	0.7323 $\pm$ 0.053	0.017 $\pm$ 0.047	8.38 $\pm$ 3.66	7.87 $\pm$ 3.19
<b>F6_L1</b>	50	0.7575 $\pm$ 0.045	0.7627 $\pm$ 0.049	0.007 $\pm$ 0.034	8.50 $\pm$ 3.46	8.08 $\pm$ 3.26
<b>F6_L2</b>	50	0.7425 $\pm$ 0.063	0.7221 $\pm$ 0.058	-0.029 $\pm$ 0.026	8.13 $\pm$ 3.31	7.67 $\pm$ 3.22



**Fig. 3.2.** Allelic Richness over eight loci for each generation and selected line independently (random breeding populations black bars, S-line light grey bars, L-line dark grey bars), calculated using a minimum sample size of N=32 (F1). Error bars represent standard deviations from averages over all loci. Allelic Richness was significantly higher in the F0 and F1 generations than the generations being bred in the lab, but no further decline over generations was observed, except for the control line, for which Allelic Richness significantly declined over generations F3-F6.

**Table 3.2.** Number of alleles observed for each locus in each generation for random breeding lines (white), S-lines (light grey) and L-lines (dark grey). Mean number of alleles over all loci and range of number of alleles per locus over all populations are given as standard deviations.

n genotyped/ n breeding	F0	F1	F2	F3					F4				
	46/90	32/50	192/250	52/52	45/50	47/50	49/50	50/50	50/50	50/50	48/50	49/50	49/50
	S1	S2	C	L1	L2	S1	S2	C	L1	L2			
Pret69	15	12	14	10	13	13	13	12	6	12	12	11	11
Pret77	9	12	12	10	8	7	10	10	8	11	8	10	9
Pr39	10	8	10	8	8	8	7	7	8	8	8	6	8
Hull70-2	20	18	21	15	15	18	16	16	16	17	14	14	16
G102	8	7	9	7	8	6	7	7	7	6	6	7	6
Pret32	16	12	15	14	13	12	11	12	13	13	10	11	12
G82	3	4	4	3	4	3	3	3	4	4	3	3	3
Pr92	7	6	6	6	6	5	5	5	6	6	5	6	6
G289	14	13	14	12	12	12	12	13	11	9	12	9	12
Over all loci	11.33	10.22	11.67	9.44	9.67	9.33	9.33	9.44	8.78	9.56	8.67	8.56	9.22
St. dev.	5.29	4.32	5.12	3.88	3.71	4.74	4.15	4.22	3.83	4.10	3.64	3.36	3.96

n genotyped/ n breeding	F5					F6					Over all populations	
	46/50	43/50	43/50	46/50	49/50	50/50	50/50	50/50	50/50	50/50	range	st. dev
	S1	S2	C	L1	L2	S1	S2	C	L1	L2		
Pret69	8	11	10	10	11	7	10	9	11	10	6-15	2.13
Pret77	7	10	7	8	9	9	8	8	7	8	7-12	1.50
Pr39	6	8	7	5	7	7	8	8	6	8	5-10	1.16
Hull70-2	14	18	15	14	14	15	14	14	14	13	13-21	2.10
G102	7	6	6	7	6	6	6	6	7	5	5-9	0.88
Pret32	13	14	10	10	12	11	11	12	11	9	9-15	1.69
G82	3	3	3	3	3	3	4	3	3	3	3-4	0.45
Pr92	5	6	5	6	5	5	6	5	6	6	5-6	0.57
G289	10	10	11	9	11	11	10	10	9	11	9-15	1.53
Over all loci	8.11	9.56	8.22	8.00	8.67	8.22	8.56	8.33	8.22	8.11		
St. dev.	3.62	4.53	3.63	3.24	3.64	3.67	3.05	3.43	3.35	3.10		

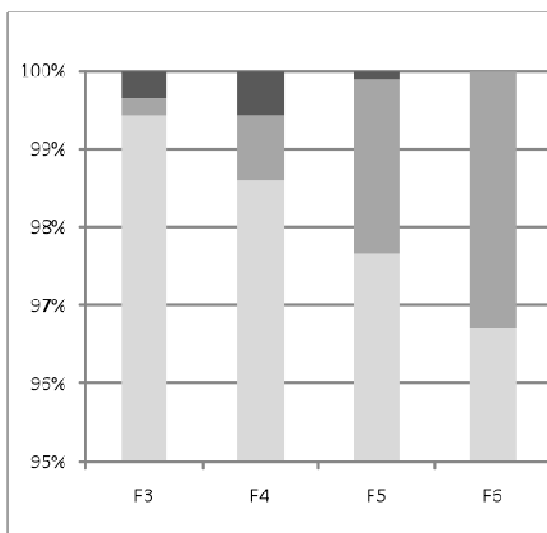
### 3.3.4 Genetic structure

#### AMOVA

A hierarchical AMOVA indicated that the majority of molecular variation was attributable to variation within lines in generations F3-F5 but that this proportion was declining over generations with a simultaneous increase in variation between lines within treatments (Fig 3.3). Total variance however increased over generations and absolute values of both variance components increased accordingly (Table 3.3). Variation between treatments was smallest and decreased over generations, both proportionally and absolute, revealing that variation was not a result of selection

regime, but rather a stochastic process resulting in increasing divergence between lines over generations.

A temporal hierarchical AMOVA of among lines ( $F_{CT}$ ), among lines between generations ( $F_{SC}$ ), and within lines within generations ( $F_{ST}$ ), revealed that the increasing variance between lines over generation was foremost the result of variation within lines within generations (98.27% of variation,  $F_{ST} = 0.0172$ ,  $p = 0.000$ ) and increasing differentiation between lines (1.41% of variation,  $F_{CT} = 0.014$ ,  $p = 0.000$ ) rather than of temporal variation within lines (0.31% of variation,  $F_{SC} = 0.0032$ ,  $p = 0.0001$ ). Full AMOVA tables per locus can be found in Appendix III.



**Table 3.3.**  $F_{CT}$ ,  $F_{SC}$ , and  $F_{ST}$  values accompanying Fig. 3.3. Significance levels are indicated as: ns (not significant), \* ( $p < 0.05$ ) and \*\*\* ( $p < 0.001$ ).

	Among treatments (F <sub>CT</sub> )	Among lines within treatment (F <sub>SC</sub> )	Within lines (F <sub>ST</sub> )
<b>F3</b>	0.00349 ns	0.00219 ns	0.00567 ***
<b>F4</b>	0.00577 *	0.00832 ***	0.01404 ***
<b>F5</b>	0.00114 ns	0.02242 ***	0.02353 ***
<b>F6</b>	-0.00476 ns	0.03293 ***	0.02833 ***

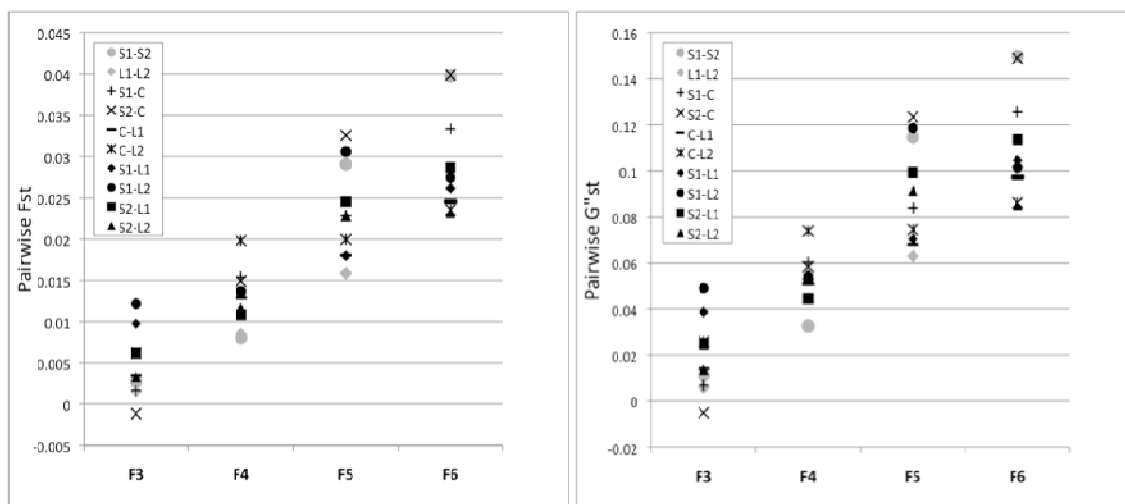
**Fig. 3.3.** Hierarchical analysis of molecular variance of eight microsatellite loci (a) among treatments (black) among lines within treatment (dark grey) within lines (light grey), showing an increasing contribution of among lines-within treatment variance at the cost of variation within lines and among treatments.

### Pairwise differentiation

Pairwise tests of differentiation revealed significant differentiation between lines at microsatellite loci, and increasing differentiation over generations. In the F3, only  $F_{ST}$  values of the S1-L1 and S1-L2 comparisons were significant after B-H corrections. All  $F_{ST}$  estimates in the F4 to F6 had  $p$ -values  $< 0.0001$  and were highly significant, although this skew in the  $p$ -value distribution did not allow SGoF+ to provide corrected  $p$ -values. Pairwise  $F_{ST}$  revealed an increase in differentiation between lines over generations but no apparent trend of greater differentiation among treatments

than within treatments was observed (Fig. 3.4). This observation was supported by an ANOVA of both  $F_{ST}$  and  $G''_{ST}$  of within treatment variation, variation among treatments and comparisons with the control line. The only significant difference observed was in the F4 among within-treatment comparisons ( $S_1-S_2$ ,  $L_1-L_2$ ) and comparisons to the control line ( $F_{2,7} = 8.661$  and  $10.439$ ,  $p = 0.013$  and  $0.008$  for  $F_{ST}$  and  $G''_{ST}$  respectively). Differentiation between lines was thus not affected by selection. This is in agreement with the pattern obtained from the AMOVA analysis that showed that little variance can be attributed to variation among treatments.

Pairwise  $G''_{ST}$  revealed a similar trend of increasing differentiation over generations and no clear pattern of increased differentiation between treatments, but the range of  $G''_{ST}$  values observed was much greater than that of  $F_{ST}$  values (per locus maximum  $G''_{ST}$  was twice the maximum  $F_{ST}$  ( $0.41$  and  $0.20$ , respectively), and over all loci this difference was more than threefold ( $0.14$  and  $0.04$ , respectively)). This can be explained by the generally high level of heterozygosity of microsatellite markers, which affects corrected  $G''_{ST}$  but not  $F_{ST}$  (Meirmans and Hedrick 2011). Whilst it is not of great interest here for investigating differentiation between lines,  $G''_{ST}$  is useful for comparing microsatellite differentiation to that of SNP markers in Chapter 4. Full tables of  $F_{ST}$  and  $G''_{ST}$  estimates per locus can be found in Appendix IV.



**Fig. 3.4.** Pairwise differentiation between selection lines estimates as  $F_{ST}$  (a) and  $G''_{ST}$  (b), showing increasing differentiation over four generations but no apparent effect of selection regime on the amount of differentiation. Grey markers indicate within-treatment comparisons, black markers comparisons between selection lines and cross and line markers comparisons between selection lines and control line.

### **3.3.5 Paternity reconstruction**

Paternity reconstruction using all ten microsatellite loci was successful and 88% ( $\pm 0.068$  st. dev.) of males could be assigned to a father in the F3-F6 generations. When correcting for missing samples in the parent generation, this proportion was even higher, at 92% ( $\pm 0.058$  st. dev.). On average, 51% of males fathered 1.7-2.7 male offspring that were retained in the next generation. Out of the wild caught males (F0) genotyped, only 9% fathered F1 males. This low value can partly be explained by missing samples (DNA was only obtained for 46 out of 90 F0 males and 32 out of 50 F1 males) but also reflects that F0 females were gravid when caught. It can therefore be expected that a substantial proportion of F1 offspring were sired by non-sampled males, as well as by new fertilisations using stored sperm. In the F2 and F3, only a low number of individuals could successfully be assigned a father as well (43-44%). Here, this is likely the result of missing samples, as only 32 out of 50 F1, and 192 out of 250 F2 males were genotyped (Table 3.4).

Using the paternity data and standard length of all males, a parent-offspring regression was performed to investigate whether father size was a predictor of male offspring size. Fig. 3.5 shows the relationship between father standard length and offspring standard length, as a deviation from population means, over all samples with fathers in generations F2-F5 and is highly significant ( $R^2 = 0.311$ ,  $F_{1,580} = 262.049$ ,  $p = 0.000$ ). Fathers in the F0 and F1 and their offspring were excluded from this analysis, because maternal effects and effects from variable rearing conditions in the wild may have affected standard length in these populations.

The slope of the regression is an estimate of half the narrow sense heritability  $h^2$ , implying a  $h^2 > 1$  when estimated over all samples and assuming no y-linked genetic variation. This estimate is no different when obtained for the control line independently (two-sample t-test,  $T_{(2),422} = 0.24$ ,  $p = 0.807$ ), but is lower and weaker for the L-lines (slope = 0.3,  $R^2 = 0.03$  and  $p = 0.016$ ) and not significant for the S-lines using this method (Table 3.5). Between generations, the slope for the regression did not differ ( $p > 0.369$ ) with exception of the F5-F6: no significant slope was observed for the parent offspring regression in this generation. Further analysis revealed that

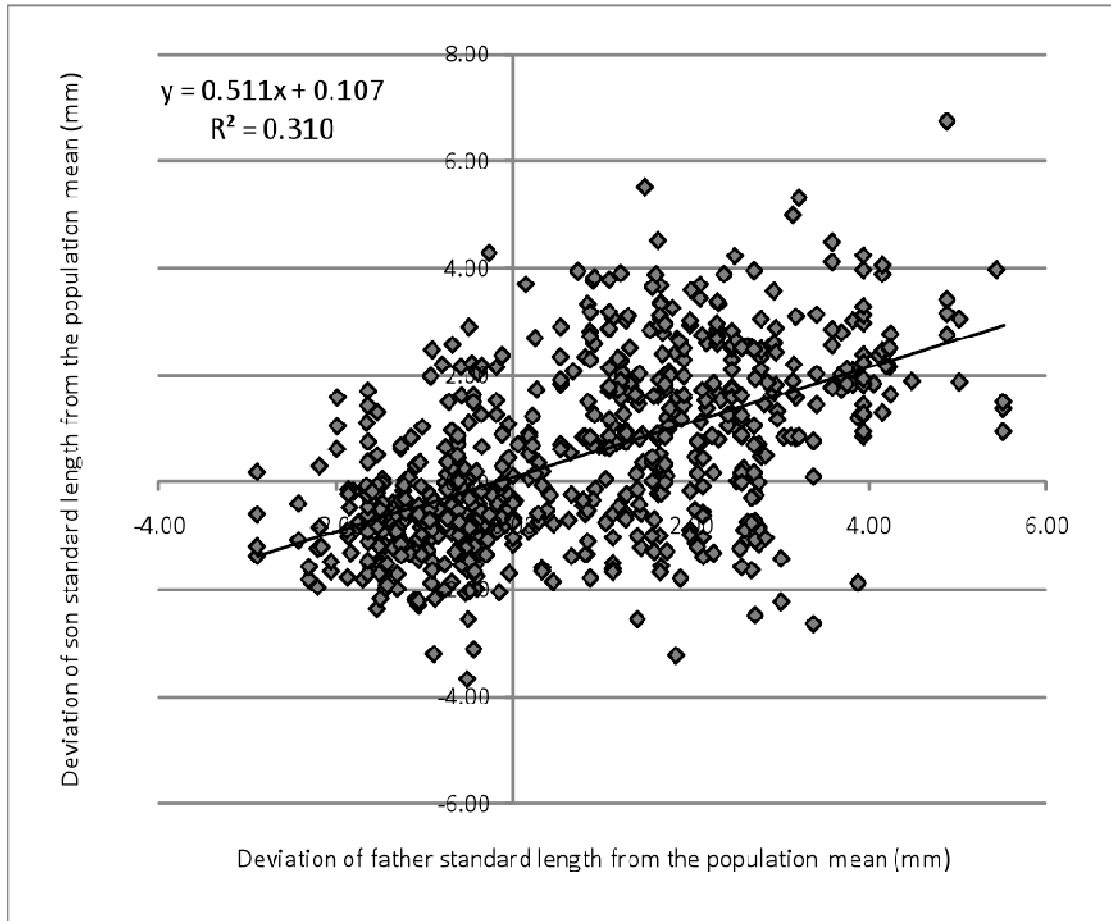
this is the result of both selection treatments, and a significant slope was present in the F6\_C ( $R^2 = 0.239$ ,  $F_{1,36} = 11.294$ ,  $p = 0.002$ , slope =  $0.489 \pm 0.146$ ). This indicates a reduced  $h^2$  of standard length as a result of selection, but not of random breeding.

**Table. 3.4.** Summary of paternity reconstruction of fish in each generation and selection line. Proportion of males fathering and mean number of offspring per father are shown for each generation, as well as the percentage of offspring that was confidently assigned (95% confidence and no mismatching loci). Based on the proportion of males fathering, the number of missing samples and the percentage of confidently assigned offspring, a final estimate of confidently assigned offspring is provided as: *No of conf. assigned offspring + (prop. of sampled males fathering x no. of missing samples x mean no. of offspring per father)*.

	F0	F1	F2	F3				
				S1	S2	C	L1	L2
N genotyped/ N breeding	<b>46/90</b>	<b>32/50</b>	<b>192/250</b>	<b>52/52</b>	<b>45/50</b>	<b>47/50</b>	<b>49/50</b>	<b>50/50</b>
Prop of sampled males fathering	0.07	0.75	0.33	0.38	0.53	0.45	0.37	0.34
Mean offspring / father	1.00	3.46	1.68	2.40	1.96	1.95	2.22	2.65
St. dev.	0.00	3.87	1.10	1.57	1.30	1.36	1.17	2.98
% confidently assigned	0.09	0.43	0.44	0.92	0.94	0.85	0.82	0.92
"" incorporating % of missing males	<b>0.20</b>	<b>0.68</b>	<b>0.57</b>	<b>0.96</b>	<b>1.04</b>	<b>0.91</b>	<b>0.83</b>	<b>0.92</b>

	F4					F5				
	S1	S2	C	L1	L2	S1	S2	C	L1	L2
N genotyped/ N breeding	<b>50/50</b>	<b>50/50</b>	<b>48/50</b>	<b>49/50</b>	<b>49/50</b>	<b>46/50</b>	<b>43/50</b>	<b>43/50</b>	<b>46/50</b>	<b>49/50</b>
Prop of sampled males fathering	0.44	0.34	0.46	0.37	0.51	1.54	0.44	0.56	0.41	0.49
Mean offspring / father	1.79	2.29	1.82	2.50	1.84	1.92	1.79	1.71	2.32	1.92
St. dev.	1.14	1.16	1.37	2.04	1.28	1.53	1.03	1.12	1.29	1.21
% confidently assigned	0.93	0.91	0.93	0.98	0.94	0.96	0.68	0.82	0.88	0.92
"" incorporating % of missing males	<b>0.93</b>	<b>0.91</b>	<b>0.97</b>	<b>1.00</b>	<b>0.96</b>	<b>1.04</b>	<b>0.79</b>	<b>0.95</b>	<b>0.96</b>	<b>0.94</b>





**Fig. 3.5.** Father-son regression of standard length over all males and all generations, showing linear relationship as:  $y = 0.511X + 0.107$  and  $R^2 = 0.310$ .

**Table. 3.5.** Relevant values from the father-son regression of standard length over all treatments and for each treatment and generation independently, showing  $R^2$  and slope of the regression, and standard error of the slope. The slope of the regression is an estimate of the narrow sense heritability.

	F2-F5 Fathers	F2 Fathers	F3 fathers	F4 Fathers	F5 fathers	Control line	L-lines	S-Lines
$R^2$	0.311	0.297	0.622	0.572	0.008	0.234	0.030	0.001
$p$	0.000	0.000	0.000	0.000	0.238	0.000	0.016	0.588
<b>N</b>	582	84	174	151	173	256	189	212
slope	0.515	0.699	0.671	0.730	-0.065	0.499	0.300	0.052
st. error	0.032	0.119	0.040	0.052	0.055	0.057	0.124	0.095

### **3.4 CONCLUSIONS AND DISCUSSION**

#### **3.4.1 The effect of selection on neutral microsatellite loci**

Here, we reported on neutral genetic diversity at eight microsatellite loci over six generations of captive breeding of guppy selection lines, and have demonstrated that three generations of divergent selection did not significantly affect neutral genetic diversity among treatments, although selection was intense (only 20% of males retained) and breeding population sizes finite and relatively small (50 males, 75 females). No deviations from HW-equilibrium were observed, and no significant inbreeding resulted from the breeding regime. No trends over generations, or significant differences between treatments, were observed for Allelic Richness,  $H_o$  or  $H_s$ , indicating that selection was not the cause of variation in genetic diversity at microsatellite loci between lines. However, Allelic Richness was significantly higher in the F0 and F1 than in later generations. The main difference between the F0 and F1 and the other generations, lies in the fact that F0 are the offspring of wild fish, and a considerable proportion of the F1 are sired by wild males through sperm storage mechanisms, as supported by the paternity analysis. The high allelic diversity seen in these generations therefore likely reflects the wild gene pool of the sampling site in the Tacarigua River more closely than the generations with only laboratory-born parents.

AMOVA in generations F3-F6 revealed genetic variation was mostly partitioned within- and among lines, the contribution of the latter proportionally increasing over generations, and among-treatment variation contributed very little to none to total variation. Therefore, the analysis confirms that drift seems the most likely explanation of the observed genetic differences between lines, rather than selection affecting microsatellite loci. Additionally, a temporal AMOVA showed that variation within lines over generations did not majorly contribute to the total variation observed, implying that over generations, selecting only 20% of individuals for breeding in each generation did not result in marked changes in neutral genetic diversity. Pairwise differentiation, as measured by  $F_{ST}$  and  $G''_{ST}$ , confirms the pattern

obtained from the AMOVA: a steady increase in differentiation between lines was observed over generations, but no pattern between treatments was detected.

### **3.4.2 The magnitude of drift through captive breeding**

Over three generations of captive breeding, divergence among lines was considerable: mean pairwise  $F_{ST}$  and  $G_{ST}$  increased on average from 0.0046 to 0.029 and variation observed between lines from 0.57% to 3.31%; a 6.3 and 5.8-fold increase respectively. The lack of concordance of this increase with the selection regime indicates that drift, rather than divergent selection caused divergence between lines at microsatellite loci.

How strong is drift here, compared to drift observed in other experiments? Simões et al. (2008a, 2008b) observed no significant  $F_{ST}$  values after three generations of laboratory breeding of *Drosophila* lines introduced from the wild, but they used much larger populations (600-1200 individuals) which obviously would have reduced drift compared to our lines (Montgomery et al. 2000).  $F_{ST}$  between lines was significant after 14 generations however (0.026-0.064) and of similar magnitude to mean pairwise  $F_{ST}$  observed in the F6 here (0.029). Also after 14 generations, Morgan et al. (2003) observed mean within-treatment  $F_{ST}$  values of 0.149 in an AMOVA of six microsatellite loci in mice selection lines, consisting of 13 breeding pairs each; 2-6 times the estimate of Samões et al. but only five times our own estimate after three generations of selection, which seems therefore reasonably comparable in magnitude. However, the design of Morgan et al. controlled for within-family matings, which maximises heterozygosity and is something we did not control for. The above two studies give an indication of the magnitude of drift observed in laboratory breeding lines, but direct comparisons are difficult, because of differences in experimental design and number of generations. As mentioned in 3.1.2, few studies have investigated and quantified the effect of selection on neutral genetic diversity in an experimental context and those who did so, reported on different numbers of generations and selection intensity. Three generations of laboratory breeding is sometimes even regarded as the baseline differentiation level (Samoies et al. 2008a) whilst here it is the endpoint of the experiments.

These examples highlight the difficulties of comparing genetic data between different experiments, and suggest even greater challenges may exist when comparing genetic data from artificial selection experiments to genetic dynamics of wild populations, in particular wild fish populations with large census sizes. However, the above also emphasises the unique nature of our study, which is, to our knowledge, the only study available to look at microsatellite diversity between selection lines over so few generations. This work has demonstrated that, even with such an intense selection regime as imposed here, neutral genetic diversity is unaffected in the short term. Therefore, the data presented here provides a useful baseline model to better understand the effects of selection in Chapter 4. The proportional change between putatively neutral loci and candidate genes not only allows for inferences on demographic events underpinning selection, but also yields information on the likelihood and magnitude of a selective response.

### **3.4.3 Paternity reconstruction**

#### **Variance in male reproductive success**

Paternity analysis indicated that the 20% of males retained in each generation were fathered by, on average, 50% of the males breeding in each preceding generation. Assuming an equal number of female as male offspring for each father, and extrapolating this figure of 50% of putative fathers being represented in 10% of the produced offspring (the 20% selected males that were genotyped in each generation), it can be reasonably concluded that all, or close to 100% of breeding males successfully fathered offspring in each generation. Therefore, sexual selection and female choice were successfully minimised by our experimental design (enforcing a female-biased sex ratio and using virgin females from a high-predation population). Although sexual selection can be a considerable factor influencing reproductive success in male guppies (Houde 1997) it thus is unlikely to have significantly reduced the genetic variance in our lines.

Since no information is available on brood sizes in our experiments, comparing these estimates to those of male reproductive success elsewhere is difficult. A crude estimate of the number of fathers per brood can be made, however, with the knowledge that around 500 fry were produced over 60 days. Assuming a gestation time of 30-40 days, this implies 1-2 broods per female, with an average number of 3.33-6.67 surviving fry per brood and an average of 1.5-3 fathers per brood. This number is comparable to the results of Evans & Magurran (2000) whose conservative estimate (non-assigned offspring was removed) for the Lower Tacarigua was 1.62 sires per brood; Becher and Magurran (2004), who observed a median of two fathers per brood and a great majority of broods having 1-3 fathering males and Neff and colleagues (Neff et al. 2008) who observed a mean of three sires per brood in high predation populations from the Caroni drainage.

#### **Reasons for unassigned paternity cases**

Paternity assignment was not 100% successful and in total in the F3-F6, 73 fish could not be assigned a father, for which several alternative explanations are possible.

Firstly, due to some undocumented mortality and DNA extraction failures, most populations had at least some missing genotypes. It is likely that a fraction of the unassigned individuals has been sired by these males.

Secondly, it could be argued that contamination between lines or samples resulted in non-assignment of individuals to their respective breeding lines. In order to test for this, all mismatching samples in the F4-F6 were analysed in Cervus against all males in the paternal generation, not only those of the own line. Only five out of 73 males were confidently assigned to a father in another line. However, each of these putative fathers was from a different line and the tanks with the putative contaminating fathers were not adjacent to the breeding tanks of the original lines. This makes contamination an unlikely cause of the low paternity assignments.

Alternatively, failure to separate males and females in time before sexual maturation could have resulted in premature matings, and sperm storage could then have

allowed fry of non-selected fathers to contribute to the next generation. Although mature males were not observed in female rearing tanks, the specific observation of low paternity in the F5\_S2 (16 non-assigned F6\_S2) combined with several large and late maturing F6\_S2 males (Fig. 2.8) raises suspicion that an unselected male may have fathered offspring in this tank. Although unfortunate, if contamination by unselected males occurred, this could only result in our estimate of the response to selection being a conservative one.

Fourth, although it is unlikely, it cannot be excluded that genotyping error resulted in mismatches between father and offspring fish. For the F0-F2, it is not possible to obtain a reliable estimate of genotyping error, due to the large number of missing genotypes in these generations. However, the 73 unassigned fish in the F4-F6 had a total of 98 mismatching genotypes with their most likely father. Out of a total of 966 genotyped fish x 10 loci, this would imply a genotyping error of 1.0%, which seems a reasonable estimate of error, whilst at the same time sufficiently low not to give rise to concerns about the reliability of the data presented here.

Finally, a small number of mismatches can probably be assigned to novel mutations of microsatellite alleles, which have been shown to explain considerable mismatches in paternity analyses elsewhere, occurring at rates as high  $3.1 \times 10^{-3}$  (Jones et al. 1999).

#### **3.4.4 Heritability of standard length**

Paternity analysis revealed a strong correlation between father and son standard length, and resulted in high estimates of heritability. When estimated for treatments independently, the regression did not yield unequivocal  $h^2$  estimates for the independent treatments. The estimate for the control line was not different from the overall estimate, but no significant slope was observed in the S-lines and the slope for the L-lines was, even though significant, much lower (0.3) and support weaker ( $R^2 = 0.03$ ). The lack of support for a relationship between father and son standard length in the selection lines, can best be explained by the fact that the sons

in the S- and L-lines represent a non-random sample of the distribution of offspring sizes, including only the sons on the negative side of the mean for the S-lines and the sons on the positive side of the mean for the L-lines. Consequentially, any relationship between offspring and father size will be much weaker than when a full spectrum of offspring sizes is present as in the control line.

When full y-linkage of standard length is assumed,  $h^2$  estimated from parent-offspring regression still is 0.515, and thus provides a minimum estimate of  $h^2$  which is strikingly similar to the estimate assuming no y-linkage in 2.4.4; 0.518. What can explain this two-fold discrepancy?

Firstly, individuals which could not be assigned a father with confidence, or for which no size data was available (i.e. fish that died before the end of the experiment and from which DNA was taken from the corps, rather than from a tissue sample combined with a size measurement) were excluded from the parent-offspring analysis, which may have affected the slope of the regression to some extent.

Secondly, the different generations yielded similar estimates of  $h^2$ , with exception of the F5-F6, where no significant slope was detected over all treatments, but a significant slope was present for the control line. This is in agreement with the results from the selection experiments, which showed no response to selection in the F6 (Fig. 2.6, 2-sample randomisation test, 10000 randomisations:  $p > 0.11$  for each of the L- and S- selected lines). A lower response to selection in the F6 is expected, as additive genetic variation is decreasing over generations and consequentially reduces the response to selection (Fig. 2.10). Consequentially,  $h^2$  estimates for the R/S regression are biased downwards by the F6. Analysis of the R/S regression without the F6 yielded a slightly higher estimate of  $h^2$  at  $2 \times 0.315 = 0.63$  ( $R^2 = 0.989$ ,  $F_{1,9} = 802.447$ ,  $p = 0.000$ , slope =  $0.315 \pm 0.011$ ). However, 0.315 is still considerably lower than 0.515 as estimated in 3.3.5.

Thirdly, and most importantly, it was shown here that the estimate from the parent-offspring regression could not be used to estimate  $h^2$  for selected lines, and the

response to selection could not be used to estimate  $h^2$  for the control line. Consequentially, the data points on which each regression is based are different: control line data is excluded from the analysis in 2.4.4 and selection line data contribute little to the estimate in 3.3.5. For this reason, it could be argued that the two estimates can be considered independent and a much higher  $h^2$  for random breeding fish than those under selection could be concluded. Reduced heritability can be expected as a result of selection, by the Bulmer effect, changes in gene frequencies at underlying loci and effects of selection on variance (Falconer and Mackay 1996; Fig. 2.10). The response to selection may therefore be a less accurate estimate of  $h^2$  in the original population than the estimate from the parent-offspring regression of the control line.

### **3.4.5 Overall conclusions**

In summary, the data presented here provides substantial support for the hypothesis that neutral genetic diversity is not significantly affected by the selection regime. In addition, these results indicate that selection and captive breeding have not resulted in significant levels of inbreeding in selection lines nor affected levels of neutral genetic variation. Rather, we have demonstrated through hierarchical AMOVA that drift between lines had a significant effect on genetic variation, but that within lines no major stochastic changes in microsatellite genetic variation occurred over generations.

Considerable discrepancies were observed between the  $h^2$  estimates from the regression in Chapter 2 and the estimate from the parent-offspring regression here. Heritability in the control line, as estimated from the parent-offspring regression (0.499-0.998), is twice that estimated in Chapter 2 (0.258-0.518). This indicates a greater reduction in additive genetic variance ( $V_A$ ) by the selection regime than initially estimated in 2.3.5. However, the microsatellite data show that this difference in  $V_A$  between selection- and control lines is not reflected in levels of microsatellite genetic diversity, as these were not different between selection and control lines. Thereby, these results provide indirect support that neutral genetic



variation is not a suitable estimator of quantitative genetic variation (Merilä and Crnokrak 2001; McKay and Latta 2002; Luttikhuisen et al. 2003; Hutchings et al. 2007). In addition, these results thus confirm that the microsatellite loci used are not affected by selection and provide a suitable baseline to compare changes at putative candidate loci to. Furthermore, the absence of significant inbreeding levels or major declines in neutral genetic variation, indicate that the breeding regime and experimental design used were successful in maintaining genetic variation over generations and genetic erosion has not likely affected the response to selection in our experiments.

4.

**GENETIC DYNAMICS OF**  
**SELECTION LINES OF *P. RETICULATA* II:**  
**IDENTIFICATION OF A RESPONSE TO**  
**SELECTION AT CANDIDATE LOCI**

## **4.1 INTRODUCTION**

The ultimate challenge for understanding evolution lies in revealing the genetic basis and pathways underpinning phenotypic variation, thereby identifying the constraints on evolution and the correlations between traits (Lewontin 1974; Stinchcombe and Hoekstra 2007). The aim of the present study is to identify genetic changes associated with phenotypic shifts in response to strong directional selection on body size in Trinidadian guppies. In previous chapters, I have demonstrated a rapid phenotypic response to selection and revealed the presence of additive genetic variation for standard length in male guppies. Furthermore, I demonstrated that neutral genetic variation measured at microsatellite loci did not change in response to selection. Here, patterns of genetic variation at specific candidate loci<sup>2</sup>, are investigated more closely to see whether a response to selection can be observed at individual loci.

### **4.1.1 A candidate gene approach to identifying functional genetic variation**

As summarised in Box 1, several different approaches are available for identifying genes and gene regions under selection in contrasting environments. Such approaches have benefitted greatly from recent developments in sequencing technology. With the candidate gene approach, knowledge on gene function, structure and chromosomal location is utilised, and the spectrum of putative gene associations with trait variation is smaller than with high-throughput sequencing approaches. In particular, the candidate gene approach maximises efficiency of use of available genomic resources and expands our understanding of additional species using relatively low sequencing effort (Piertney and Webster 2010).

---

<sup>2</sup> A candidate gene is defined by (Fitzpatrick et al. 2005) as: “any gene that has been identified in one organism and is predicted to have a similar phenotype in another”. De-Santis and Jerry (2007) make a further distinction between candidate genes and major genes: a candidate gene being a gene of which the physiological function is known and its direct effect on the expression of a trait quantifiable. A major gene is a gene or marker of which the exact physiological role is uncertain but a strong association of the gene with trait variation is observed. Both major genes and candidates are included in the present study when candidate genes are being referred to.

For example, comparison of a QTL region to which coat colour polymorphism in soay sheep (*Ovis aries*) was mapped with the cattle genome allowed the identification of a candidate gene for coat colour (Beraldi et al. 2006; Gratten et al. 2007). Subsequent work revealed a significant association with a non-synonymous substitution (single nucleotide polymorphism SNP) at this locus likely to be causal in coat colour variation (Gratten et al. 2007).

Owing to their role in immune response and pathogen resistance, genes from the major histocompatibility complex (MHC) have been repeatedly characterised and demonstrated to be targets of selection. Successful associations with e.g. parasite load and reproductive success have been identified in a range of non-model species (Bernatchez and Landry 2003; Lukas et al. 2004; Oliver et al. 2009), including guppies (Sato et al. 1995; van Oosterhout et al. 2006b; Fraser et al. 2009, 2010).

Fidler et al. (2007) sequenced a dopamine receptor to investigate its role in novelty seeking behaviour in great tits (*Parus major*). The birds were the result of four generations of selective breeding on exploratory behaviour and polymorphisms in the candidate gene were associated with the different selection regimes. In a subsequent study, they identified variation in behavioural phenotypes in the wild source population and confirmed one of the genetic associations (Korsten et al. 2010), providing further support for a functional role of the gene and associated polymorphism in behavioural phenotypes. These examples highlight the usefulness of selection experiments for investigating the genetic basis of ecologically relevant traits and Fidler et al. (2007) is to my knowledge the only study doing so for an experimental design comparable to our own.

In previous chapters, I have discussed the work on the *Pan I* locus in cod (*Gadus morhua*), which provides an informative case study for the use of candidate genes in identifying the genetic response to selective pressures in the wild (Pogson 2001; Jónsdóttir et al. 2008; Árnason et al. 2009; Jakobsdóttir et al. 2011).

Genes involved in growth have been a particular target of research in fishes, as a result of the rapid development of the aquaculture industry and the associated increase in interest in marker-assisted breeding techniques (Jackson et al. 1998; De-Santis and Jerry 2007). With rising concerns for genetic changes in growth and reproduction by selective fishing (see 1.2.3 and 1.2.4), this interest has expanded beyond the aquaculture industry and considerable effort is targeted at characterising growth- and reproduction related genes for exploited fish species. For example, Tao and Boulding (2003) identified SNPs in ten growth-related candidate genes in Arctic charr (*Salvelinus alpinus*) and observed an association between early growth and one of these markers. Hemmer-Hansen et al. (2011) used a candidate gene approach to scan 18 growth- and reproduction-related genes for polymorphism in cod and identified over 30 observed polymorphisms between five cod populations. Such efforts deliver useful tools for analysing both contemporary and historic samples and yield empirical opportunities to assess fisheries induced evolution (FIE).

#### **BOX 4. CANDIDATE GENES FOR GROWTH AND REPRODUCTION**

For investigating the genetic basis of body size, genes of several different pathways are of particular interest, which are summarised briefly below.

##### **Growth axis**

The growth axis, or somatotropic axis, is the main pathway regulating somatic growth and consists of growth hormone (GH) in the pituitary gland, which stimulates the production of insulin-growth factors (IGF) that stimulate metabolic processes and tissue growth (Moriyama et al. 2000, Florini et al. 1996). The pathway consists of GH and IGF-I and II, and associated releasing hormones (GHRH), inhibitors (GHIH or somatostatin), receptors (GHR, IGFR), binding proteins (IGFBP) and carriers (GHC). All these genes have been associated with variations in weight and size of various fish species (e.g. Tao and Boulding 2003; Ryyanen and Primmer 2004; Almuly et al. 2005). The majority of these associations however are concerned with intronic point mutations or microsatellites, and do not represent causal mutations explaining variations in growth and size.

##### **Transforming growth factors**

Myogenic regulatory factors (MRF; myogenin, the MyoD complex, myf-5 and myf-6) and growth factors (GF) are part of the transforming growth factor superfamily and mediate the hormone signal sent by the pituitary gland and structural growth at the muscle tissue level (Atchley et al. 1994). Myostatin (MSTN) is a specific growth factor and inhibition of this gene has been reported to improve growth in mice with up to 30% (McPherron et al. 1997). Since then, polymorphisms in MSTN have been associated with growth differences in a range of species. In mammals, MSTN is expressed exclusively in skeletal muscles and inhibits muscle development but in fish it has been reported in other tissues as well, including gills and gonadal tissue, suggesting a more general role of MSTN in physiological development in fish.

#### **BOX 4. (CONTINUED)**

##### **Maturation; hypothalamic-pituitary-gonadal (HPG) axis**

Genes involved in maturation and reproduction are of interest in light of FIE because of the proposed selective effects of fishing on maturation and because of the trade-offs between growth and reproduction-related traits and strong linkage between the two. Major candidates involved in the maturation process are aromatase, which controls the androgen/oestrogen ratio (Kroon et al. 2005), and the gonadotropins follicle-stimulating hormone and luteinizing hormone (FSH and LH), which regulate egg/ sperm production, with associated releasing- and receptor genes. Other genes connected to the HPG axis and the maturation process are leptin, melanocortin-4 receptor and steroidogenic factor genes.

##### **Food intake and metabolism**

Leptin, which is thought to be also minorly involved in GH regulation (Heiman et al 1998), has been associated with food intake and metabolism in fishes (Johnson et al. 2000).

Grehlin is an important hormone affecting food intake in fish, and is also involved in regulation of pituitary hormones like GH and prolactin (Unniappan and Peter 2005).

The melanocortin-4 receptor (MC4-R) gene, is a candidate gene that has been shown to affect body size through both regulation of the energy budget affecting food intake and through leptin regulation and gonadotropin production (Martin et al. 2006), affecting the onset of maturation. In platyfish *Xiphophorus* spp. this gene is found on the sex-chromosome and copy number variation in alleles is a major candidate for the P-locus, which is known to underlie different male size morphs by determining the timing of maturation (Kallmann 1983; Lampert et al. 2010).

##### **Other putative candidates**

Jun transcription factors are genes involved in regulating gene activity following the primary growth factor response (Schlingensiepen et al. 1994).

Parvalbumins (PVALB-I and II) are involved in fibre relaxation and calcium binding in muscle tissue and a microsatellite polymorphisms within the PVALB-I gene has been associated with weight in seabass (*Lates calcarifer*) (Xu et al. 2006).

Pituitary adenylate cyclase-activating polypeptide (PACAP) is found on the same mRNA precursor as GHRH and stimulates the release of GH, prolactin, adrenocorticotrophic hormone (ACTH) and luteinizing hormone (LH) (Miyata et al, 1989). A SNP in this gene (intron) has been associated with growth in charr (Tao and Boulding 2003).

Prolactin is a hormone for which a range of effects has been described in vertebrates, including effects on growth and development (Bole-Feysot et al. 1998). Perhaps most known in fishes is its role in osmoregulation and freshwater adaptation (Manzon et al. 2002; Blel et al. 2010). Its inhibitor, dopamine, is furthermore involved in light adaptations (Witkovsky 2004)

Somatolactin is a pituitary hormone related to growth hormone and prolactin and involved in gonadal maturation and spawning in fish (Ono et al. 1990; Rand-Weaver et al. 1992).

PIT-1 is a pituitary-specific activator of GH and thus involved in regulating somatic growth (Tuggle and Trenkle, 1996).

Furthermore, genes involved in a general stress response, such as heat shock proteins (HSP) may be expected to respond to adaptations to the experimental laboratory environment (Parsell and Lindquist 1993).

#### **4.1.2 A context for the identification of polymorphisms and genetic association studies**

Despite the increasing availability of gene associations with phenotypic variation in ecologically relevant traits, it is important to realise that the identification of polymorphisms and association studies are only a first step towards understanding the genetic basis of the traits in question. How polymorphisms translate into variation at the phenotypic level, by altering gene transcription and physiological processes, is an important sequel to candidate gene analysis (Vasemägi and Primmer 2005; Stinchcombe and Hoekstra 2007). However, how such functional genomic mechanisms occur is still rarely explained, given the large number of studies that report traits associated with synonymous or intronic polymorphisms (e.g. in DeSantis & Jerry 2007).

#### **4.1.3 A candidate gene approach to investigating the response to selection on body size**

The present study utilises candidate genes for growth and body size, in an attempt to characterise some of the variation that may underlie the response to selection observed in our experiments. The choice of a candidate gene approach, rather than a genome-wide search for putative candidates, was fuelled by feasibility in light of limitations to resources and time, but especially by the fact that for guppies, a plethora of genetic resources is available and a candidate gene approach makes optimal use of these resources.

Candidate genes for investigating a genetic response to selection in the present study were obtained from two different sources. Firstly, six SNP markers and their flanking regions that were identified to be in tight linkage with a major QTL for male body length on the sex chromosome in the linkage map of colour and male size traits by Tripathi et al. (2009a). Secondly, genes for which existing knowledge from other species suggests a role in growth and size determination (Box. 4) and for which sequences of *P. reticulata* were available on GenBank. These markers were used for the identification of polymorphisms and a range of genetic analyses was used to identify signatures of selection at individual marker loci.

**Table 4.1** Candidate genes for standard length in male *P. reticulata* investigated in the present study. The top box lists markers obtained from Tripathi et al. 2009 (M9-M1078); the bottom box candidate genes obtained from a *P. reticulata*-specific query on GenBank are listed. Query term, putative functions of obtained markers, GenBank association number and DNA type of sequences are given. If SNPs were successfully developed for the marker/ gene in question, SNP marker names are provided in the final column. Markers named \*\_Dreyer are the same SNP markers as used by Tripathi et al. 2009a for that specific marker.

Query	<i>P. reticulata</i> sequence obtained / putative function	GenBank No.	Sequence	SNP Marker name(s)
M009	Rab interacting lysosomal protein-like 1	ES375672	cDNA	M9_indel, M9_403
M030	Fructose-1,6-bisphosphatase 1	ES382008	cDNA	M30_Dreyer
M061	Ectodermal-neural cortex 1 ENC1	ES379217	cDNA	---
M155	Splicing factor, arginine/serine-rich	ES385036	cDNA	---
M380	Medaka Chrom. 15, in scaffold91_contig45820	FH889280	BAC genomic	---
M987	---	FH893254	BAC genomic	M987_indel
M1046	Melanocortin 4 Receptor	FI236224	mRNA, partial cds	M1046_Dreyer, M1046_2
M1078	Insulin growth factor 1	FI236241	Intron	---
Steroidogenic factor	SF1 nuclear receptor 5A1 steroidogenic factor	FI236230	mRNA, partial cds	SF1
Insulin growth factor	Insulin growth factor 2	DQ337476	mRNA, complete cds	---
IGF-Binding protein	IGF-Binding Protein 1	FI236240	mRNA, partial cds	---
Growth hormone	Growth hormone	U63805	partial cds	GH1, GH2_60, GH2_74, GH2_165, GH2_211
---	putative GH-regulated TBC protein	ES386819	cDNA	TBC1
Heat Shock Protein	similar to Heat-shock protein beta-1 (Hspb1)	ES385844	cDNA	---
Myostatin	similar to myostatin; isoform II	ES380902	cDNA	Myostatin
Pit-I	similar to PIT54 (Gallus gallus)	ES377286	cDNA	---
Transcription factor	Transcription factor JunB	EF408832	cDNA, partial cds	---
Prolactin	similar to prolactin 1	ES374674	cDNA	Prolactin_1, Prolactin_2



## **4.2 MATERIALS AND METHODS**

### **4.2.1 DNA Sampling and extraction**

Sample preparation and DNA extraction protocols are provided in 2.3 and 3.2.

### **4.2.2 Candidate marker choice**

#### **QTL-mapped markers**

Six markers were screened that were previously found to be in close linkage with a major QTL for male standard length (M0009, M0030, M0061, M0155 M0380 and M0987; Tripathi et al. 2009a). Two further SNP markers from this study were used; M1046, the partial coding sequence (cds) of the melanocortin 4 receptor (MC4R) and M1078, the intronic sequence of insulin-like growth factor 1 (IGF1).

#### **Growth-related genes**

A list of 31 genes and gene families was assembled that can reasonably be expected to be involved in growth- and maturation-related pathways, based on knowledge of gene function in other (fish) species. A search was performed in the GenBank NCBI nucleotide collection (nr/nt) for published sequences of *P. reticulata* for any of these genes (Table 4.1). Sequences were aligned with sequences from other species using BLASTn (Altschul et al. 1997) to identify conserved regions. Primers were designed within both conserved and non-conserved regions using Primer3 (Rozen and Skaletsky 2000). Since most sequences were cDNA sequences and no information was available on the position of intron/exon boundaries, primers were designed for products with a maximum size of around 500 bp. The presence of an intron within the amplification region would thus not necessarily result in lengthy fragments, which could reduce amplification success and be difficult to sequence. A total of 17 primer pairs were designed for ten genes (Table 4.1).

#### **Microsatellite marker Pr39**

Microsatellite locus Pr39 (Becher et al. 2002) was included in the candidate gene analyses. This locus has been used by others for population genetic studies in the

wild (van Oosterhout et al. 2006a; Barson et al. 2009; Suk & Neff 2009) but in our experiments, putative linkage of this marker to a locus under selection in L-lines was suggested by HW-deviations of Pr39 observed in L-lines only (see 3.3.2).

#### **4.2.3 Sequencing reactions and PCR protocols**

For the QTL-linked markers, the same primers were used for amplification as those used by Tripathi et al. (2009a), except for M1078, for which a new primer pair was designed using Primer3 (Rozen and Skaletsky 2000).

All QTL map primers amplified at 56°C annealing temperature (AT) and PCR products were obtained using a 10µl reaction (4.4µl H<sub>2</sub>O, 1.5µl MgCl<sub>2</sub>, 1µl colourless GoTaq® Flexi buffer, 1µl 10mM dNTPs, 0.5µl F primer, 0.5µl R primer and 1µl DNA, GoTaq® Flexi DNA polymerase (PROMEGA). PCR products were obtained using a Tetrad2 Peltier thermal cycler (BIO-RAD) using the following program: 95°C 5m; 95°C 1m, 56°C 1m, 72°C 1m for 35 cycles, 72°C 10m. Unless stated otherwise, all primers were at 10mM concentrations and product amplification was verified on a 1% agarose gel.

The amplification success of primers designed from sequences obtained from GenBank was tested using a gradient PCR and a 10µl reaction (4.4µl H<sub>2</sub>O, 1.5µl MgCl<sub>2</sub>, 1µl colourless GoTaq® Flexi buffer, 1µl 10mM dNTPs, 0.5µl F primer, 0.5µl R primer, 0.1µl DNA, GoTaq® Flexi DNA polymerase (PROMEGA), on a thermal cycler using the following program: 95°C 5m; 95°C 1m, AT 1m, 72°C 1m for 35 cycles, 72°C 10m).

PCRs for sequencing were conducted in 20µl reactions (8.8µl H<sub>2</sub>O, 3µl MgCl<sub>2</sub>, 2µl colourless GoTaq® Flexi buffer, 2µl 10mM dNTPs, 1µl F primer, 1µl R primer, 0.2µl DNA, GoTaq® Flexi DNA polymerase (PROMEGA) for: 95°C 5m; 95°C 1m, AT 1m, 72°C 1m for 35 cycles, 72°C 10m. Successful amplification was verified on 1% agarose gel and products were cleaned using 0.1µl Exonuclease I, 0.2µl Thermo sensitive alkaline phosphate (TSAP) and 0.7µl H<sub>2</sub>O per reaction in a thermal cycler for 37°C 15m, 74°C

15m. All sequencing reactions were performed by MACROGEN Korea (<http://www.macrogen.com>).

#### **4.2.4 SNP discovery**

In order to screen for polymorphisms of interest within markers, amplicons derived from PCR with eight primer pairs obtained from Tripathi et al. (2009a) were sequenced for the six smallest and six largest F3 individuals. Subsequently, marker M987, for which an insert was observed that was present only in S1 individuals of the first screening, was sequenced for the ten smallest and largest individuals in the F3 and F4 generation. Four further markers showing putative segregating polymorphisms between treatments (M9, M380, M1046, M30) were sequenced for the eight smallest/ largest individuals respectively of the S1 and S2 and L1 and L2 lines in the F4 generation.

Fourteen successfully amplifying primer pairs (Table 4.1) for the growth-related genes obtained from GenBank and M1078, the IGF1 intron obtained from Tripathi et al. (2009a) (obtained with new primer pairs), were sequenced for two F6 individuals from each selection line (the smallest two individuals of the S1 and S2 and the largest two individuals of the L1 and L2). In addition, a sample consisting of the pooled DNA of the ten smallest/ largest individuals of each line was sequenced for each primer pair.

Sequences were aligned using Geneious Pro 4.8.5 and manually inspected for polymorphisms. Only polymorphisms present in both directions of at least two individuals in high-quality sequences were considered reliable. In total, in addition to the two insertions/ deletions, 14 SNPs of interest were detected within eight markers (Table 4.1) and sent to KBioscience (<http://www.kbioscience.co.uk>) for genotyping.

#### **Candidate marker genotyping**

All breeding males from generation F0, F2, F5 and the males used for estimating maturation age in the F6 were genotyped for all candidate markers, thus providing marker frequencies in pre- and post selection samples. Insertions/ deletions (indels)

were observed in M9 and M987, which were incorporated in the microsatellite 58°C multiplex panel, using new primer pairs designed in Primer3 to create appropriate product sizes. For these markers, and microsatellite marker Pr39, genotypes were obtained for additional populations (all generations for Pr39, all but the F1 for M987 and all but the F1 and F3 for M9). Where appropriate, these data are included in the analysis. All primer sequences and SNP markers with flanking regions, used for genotyping assays by KBioscience, are listed in Appendix I.

#### **4.2.5 Genetic diversity and Linkage disequilibrium**

Using GenoDive (Meirmans and van Tienderen 2004), the following parameters were obtained per locus for each selection line and for each generation independently: observed heterozygosity  $H_O$ , expected heterozygosity  $H_S$ , and inbreeding coefficient  $G_{IS}$  (Nei 1987). Because for the identification of putative candidates for selection, the signal of each marker needs to be evaluated independently, summary statistics over all loci were not calculated. Instead, only values for those loci showing deviations from predicted values, or interesting patterns between lines or treatments are shown.

Genotypic linkage disequilibrium (LD) was tested for using Genepop's log likelihood ratio statistic for each pair of loci in each population. As with neutral microsatellite loci, LD was tested for in the wild-caught (F0) fish only because small breeding groups and selection may affect the non-random inheritance of loci, even when not physically linked.

#### **4.2.6 Rationale underlying candidate marker analysis**

For the purpose of detecting loci under selection, several different methods can be used, based on different properties of loci. A multi-disciplinary approach to detecting selection is important, since selection may act upon different levels, i.e. Hardy-Weinberg equilibrium, heterozygosity or allele frequencies. Marker loci may reveal a signal at one level, but evidence of selection may not necessarily be detectable at another level. For example, a single allele may show an association with body size whilst heterozygosity levels remain unchanged. Outlier analysis, based on  $F_{ST}$  or

heterozygosity, would not detect such a signal. Alternatively, heterozygosity may be under selection, which is not necessarily revealed by analysis of allele frequencies (Schlötterer 2003; Martínez et al. 2011).

Furthermore, a distinction can be made between associations of population gene frequencies and selection treatment, or associations of individual genotypes and standard length. Both can be indicative for selection at a candidate locus, but the underlying tests are different and provide support for a role of the candidate locus in the response to selection, or in the genetic basis to standard length respectively. These are two tightly linked concepts, but intrinsically different, because an association of a certain allele with standard length can be expected to be the same in all populations, whilst outlier- and frequency-based analyses are expected to result in differences between treatments for successful candidate loci.

Following Vasemagi and Primmer (2005), the analyses in this chapter are classified into four categories:

**1) Single-locus tests for neutrality - deviations from Hardy-Weinberg equilibrium.**

- a. Despite suffering from low power (Lewontin & Cockerham 1959) and potentially being biased by genotyping errors (Hosking et al. 2004), deviations from HW-equilibrium, or a significant inbreeding coefficient  $F_{IS}$ , indicates higher/ lower frequencies of certain genotypes than expected under random mating and can thus be indicative of selection (Watterson 1977, 1978; Lachance 2009). Here, deviations from HW-equilibrium were estimated for each locus in each population and for each generation independently, using a Hardy-Weinberg probability test for 10.000 batches and 10.000 iterations per batch in Genepop version 4.0.10 (Raymond and Rousset 1995; Rousset 2008). All p-values were subsequently corrected for multiple tests using Benjamini-Hochberg corrections in the program SGoF+ (Carvajal-Rodríguez et al. 2009).
- b. Both drift and selection can result in significant differences in allele frequencies at candidate loci between populations. However, when these

differences are consistent between treatments, this may indicate selection for certain alleles in different selection regimes and associations with different phenotypes (Fidler et al. 2007; Jakobsdóttir et al. 2011). Using Genepop version 4.0.10 (Raymond and Rousset 1995; Rousset 2008) a log-likelihood based exact test with 10.000 batches and 10.000 iterations per batch was used to estimate genotypic differentiation at each candidate locus between all pairs of populations in generation F5 and F6. Results were corrected for multiple tests using Benjamini-Hochberg corrections in SGoF+.

2) **AMOVA.** The hierarchical partitioning of molecular variation between and within treatments can be indicative of selection, when variation between treatments is greater than between lines within treatments. AMOVAs were carried out using Arlequin version 3.5.1.2 (Excoffier et al. 2005) for generation F5 and F6 independently using the following hierarchical levels: (1) within lines, (2) within treatments-between replicate lines, and (3) between treatments. These hierarchical levels are equivalent to  $F_{ST}$ ,  $F_{SC}$  and  $F_{CT}$  respectively. Additionally, the same hierarchical analysis was performed on both generations combined.

3) **Outlier analyses; multiple-marker-based neutrality tests**

a. Pairwise differentiation, as measured by  $F_{ST}$  and  $G''_{ST}$ , can be compared between neutral and putatively selected loci. Higher/lower divergence than expected under neutrality can be indicative of selection at outlier loci. Several widely used methods incorporate this principle for detecting selection (see above). However, these methods are based on  $F_{ST}$  and heterozygosity, which may differ between loci depending on the number of alleles and not necessarily only as a result of selection (Meirmans and Hedrick 2011). Therefore, standardised pairwise differentiation, as estimated by  $G''_{st}$ , is used here as well to provide an additional measure of differentiation useful for comparing different types of markers (i.e. microsatellites and SNPs). However, due to the large number of tests involved, this method suffers from low power and therefore a lack of significance of outlier loci may be expected.

- b. The *fdist* outlier method (Beaumont and Nichols 1996) compares a plot of heterozygosity and  $F_{ST}$  of marker loci to a simulated neutral distribution of these parameters. Based on the principle of increased genetic differentiation between loci under divergent selection, loci with greater differentiation than the computed neutral space are candidates for divergent selection, and those falling below the neutral distribution for balancing selection. Here, LOSITAN – Selection Workbench (Antao et al. 2008) was used to estimate outlier loci from the entire set of 25 markers (microsatellites and candidate loci), using 10.000 simulations, an infinite alleles model and a forced neutral mean  $F_{ST}$ . Additional simulations were run in cases where 10.000 simulations were inconclusive in determining whether a locus was under selection or not; i.e. when they were positioned on the neutral-selection interface in the plot. The analyses were performed on the entire dataset of microsatellite- and candidate loci between all pairs of populations, between all treatment pairs and over all populations in the F5 and F6 separately, as well in the F5 and F6 combined.
- c. The third outlier method used is the heterozygosity-based selective sweep method (Schlötterer 2002). This method is based on elevated rates of loss of heterozygosity of loci under selection, compared to populations where the locus is not under selection. It has been demonstrated (Schlötterer 2002; Kauer et al. 2003) that the natural logarithm of the ratio of expected heterozygosity in two populations ( $\ln RH$ ) follows a normal distribution and can be expected, after normalisation to mean=0 and SD=1, to fall between -1.96 and 1.96. Outlier loci from this distribution are candidates for selection.  $\ln RH$  was calculated here for each locus (both microsatellites and candidate loci) and each population pair within generation F5 and F6 (all within- and between-treatment comparisons), and normalised over all loci per population pair. Subsequently, outlier values were detected manually from the obtained distribution.
- 4) **Association analysis.** Using the online program SNPStats (Solé et al. 2006), associations between male standard length and genotype were tested for each candidate locus in the F5 and F6 independently, in all control

populations combined (F0, F2, F5, F6) and in a global test over all populations and generations F0, F2, F5 and F6. Each analysis was corrected for multiple testing using Benjamini-Hochberg corrections in SGoF+. The most likely model of inheritance was inferred based on the lowest  $p$ -value and lowest value for the Akaike Information Criterion (AIC) and congruence between results of different analyses. Locus M987 and Pr39, which have more than two alleles, were converted to bi-allelic loci based on putative associated alleles from other analyses (see section 4.3.3 below): the absence/ presence of an insert for M987, and the absence/ presence of allele "174" for Pr39. Five different inheritance models were considered: co-dominance, dominance, over-dominance, recessive and a log-additive model. Interactions between loci were not considered.



## **4.3 RESULTS**

### **4.3.1 SNP discovery and genotyping success rate**

#### **QTL-linked markers**

All eight QTL-linked markers amplified successfully in our selection lines (using new primers for M1078). From the SNPs described in Tripathi et al. (2009a), putative segregating SNPs were observed only in M9, M30, M380, M987 and M1046. Additionally, M987 and M9 both contained an indel that appeared to have different frequencies in different lines, based on the sequencing of the smallest and largest F3 and F4 individuals. For M987, four alleles were present: one without an insert, and three with an insert, each one basepair different in size. In addition to the two indels, two of the originally designed SNPs (M30\_Dreyer, M1046\_Dreyer) and two additional SNPs (M9\_403, M1046\_2) were ultimately selected for large-scale genotyping.

#### **Candidate genes**

Sequence hits for specific gene queries for *P. reticulata* were found for ten genes (Table 4.1) for which a total of 16 primer pairs were designed. From these, three did not amplify (primers for partial sequences of GH, myostatin and HSP), leaving 13 successful amplifying primer pairs. Two of these sequences were of insufficient quality to allow further analysis (IGF-II and a partial sequence of TBC1). A considerable proportion of the other sequences was also difficult to read and failed to provide data of sufficient quality to reliably mine for SNPs in at least one sequencing direction (5' → 3' or 3' → 5'). This is probably a result of genomic DNA being sequenced in a single reaction, rather than as isolated alleles through cloning. Any size variation between alleles (e.g. due to indels in intronic sequences) would have resulted in inconsistencies between base calling peaks and unreadable fragments in one direction of the sequence. Although this limited the data available for SNP mining, we judged this method preferable in terms of speed and cost. From the readable sequences, six were monomorphic (IGF-I, IGFBP, the second part of HSP, JunB and both parts of Pit54). The remaining five sequences yielded ten

putative SNPs: five in GH, two in prolactin, and one each in myostatin, TBC1 and SF1. Together with the four SNPs from the QTL-linked markers, 15 putative candidate SNPs were thus genotyped in a total of eight genes for all F0, F2, F5 and F6 samples.

**Table. 4.2.** Allele frequencies for all candidate markers showing frequencies of allele-specific nucleotides for SNP markers and number of basepairs for indels and microsatellite Pr39.

Locus	Allele	F0	F2	F5_S1	F5_S2	F5_C	F5_L1	F5_L2	F6_S1	F6_S2	F6_C	F6_L1	F6_L2
M9_403	A	27.91	29.95	25.56	38.10	24.42	8.89	12.24	26.53	41.84	40.00	14.00	15.00
	G	72.09	70.05	74.44	61.90	75.58	91.11	87.76	73.47	58.16	60.00	86.00	85.00
M1046_Dreye	T	36.36	30.11	31.11	39.29	34.88	22.73	36.96	32.65	43.00	36.00	30.61	28.00
	C	63.64	69.89	68.89	60.71	65.12	77.27	63.04	67.35	57.00	64.00	69.39	72.00
M1046_2	A	19.32	14.17	17.39	13.41	9.52	14.13	18.37	23.00	17.00	13.00	18.37	17.00
	G	80.68	85.83	82.61	86.59	90.48	85.87	81.63	77.00	83.00	87.00	81.63	83.00
M30_Dreyer	T	62.50	52.67	50.00	29.76	54.76	63.04	66.67	54.08	26.53	73.00	71.43	69.00
	C	37.50	47.33	50.00	70.24	45.24	36.96	33.33	45.92	73.47	27.00	28.57	31.00
GH1	T	63.95	67.96	63.04	58.33	53.85	65.12	68.37	73.00	51.00	55.32	65.31	66.00
	G	36.05	32.04	36.96	41.67	46.15	34.88	31.63	27.00	49.00	44.68	34.69	34.00
GH2_60	A	61.11	42.20	56.82	65.00	54.76	43.18	50.00	44.90	63.00	58.16	40.00	38.00
	G	38.89	57.80	43.18	35.00	45.24	56.82	50.00	55.10	37.00	41.84	60.00	62.00
GH2_74	G	71.59	83.69	77.17	72.62	84.52	84.44	78.72	77.00	79.00	82.29	84.00	86.00
	C	28.41	16.31	22.83	27.38	15.48	15.56	21.28	23.00	21.00	17.71	16.00	14.00
GH2_165	T	41.67	61.29	54.35	40.24	61.63	61.11	53.13	60.42	34.00	61.00	65.00	61.00
	G	58.33	38.71	45.65	59.76	38.37	38.89	46.88	39.58	66.00	39.00	35.00	39.00
GH2_211	A	30.68	22.99	24.44	35.37	29.76	21.59	26.53	21.88	42.86	29.00	21.74	30.00
	C	69.32	77.01	75.56	64.63	70.24	78.41	73.47	78.13	57.14	71.00	78.26	70.00
Myostatin	A	84.52	75.41	73.91	79.27	71.43	72.22	65.63	75.51	69.39	79.00	75.51	72.00
	C	15.48	24.59	26.09	20.73	28.57	27.78	34.38	24.49	30.61	21.00	24.49	28.00
Prolactin_1	A	41.11	35.90	28.89	35.71	40.48	52.22	48.96	28.00	30.00	52.04	63.00	46.00
	G	58.89	64.10	71.11	64.29	59.52	47.78	51.04	72.00	70.00	47.96	37.00	54.00
Prolactin_2	A	42.39	50.53	44.19	44.05	41.67	54.35	59.18	54.00	44.90	57.45	66.00	59.18
	G	57.61	49.47	55.81	55.95	58.33	45.65	40.82	46.00	55.10	42.55	34.00	40.82
TBC_1	T	46.67	48.14	42.05	50.00	65.12	42.39	43.75	50.00	52.00	70.41	41.00	34.00
	G	53.33	51.86	57.95	50.00	34.88	57.61	56.25	50.00	48.00	29.59	59.00	66.00
SF1	A	88.64	85.26	85.23	95.24	80.49	78.26	83.67	79.17	94.00	88.00	77.00	87.00
	T	11.36	14.74	14.77	4.76	19.51	21.74	16.33	20.83	6.00	12.00	23.00	13.00
M9_indel	270	85.56	85.05	84.78	81.40	88.37	92.39	94.90	86.00	97.00	88.00	98.00	92.00
	273	14.44	14.95	15.22	18.60	11.63	7.61	5.10	14.00	3.00	12.00	2.00	8.00
M987	307	14.44	29.08	18.48	30.23	39.53	46.74	50.00	20.00	44.00	23.00	55.00	49.00
	316	6.67	1.90	2.17	1.16		7.61	8.16				5.00	2.00
	317	77.78	60.05	78.26	53.49	53.49	43.48	34.69	74.00	35.00	71.00	39.00	44.00
	318	1.11	8.97	1.09	15.12	6.98	2.17	7.14	6.00	21.00	6.00	1.00	5.00
Pr39	168	1.09	0.53										
	174	2.17	7.67	9.78	11.63	23.26	39.13	37.76	8.00	2.00	12.00	43.00	36.00
	176	5.43	5.56	1.09	3.49	1.16			5.00	2.00	7.00		1.00
	178	54.35	42.86	43.48	58.14	46.51	26.09	36.73	34.00	66.00	51.00	29.00	41.00
	180	7.61	11.38	23.91	9.30	15.12	8.70	11.22	34.00	14.00	6.00	6.00	8.00
	182	13.04	16.93	7.61	4.65	4.65	20.65	6.12	8.00	3.00	5.00	11.00	3.00
	184		0.26										
	190	1.09											
	202	7.61	8.99	14.13	4.65	6.98	5.43	5.10	10.00	1.00	14.00	10.00	5.00
	204	5.43	3.97		6.98	2.33			1.02	1.00	11.00	4.00	2.00
	218	2.17	1.85		1.16				2.04		1.00	1.00	4.00

## Genotyping success

Genotyping success rates were high, with only 182 out of 9971 genotypes missing (1.83% failure). With the exception of SF1 which had a minor allele frequencies of 4.8% in S2\_F5, no minor allele frequencies < 5% were observed. Allele frequencies per population for all candidate loci are given in Table 4.2

### 4.3.2 Genetic diversity and linkage disequilibrium

#### Linkage disequilibrium

Out of 17 loci compared, eight combinations revealed significant linkage disequilibrium after B-H corrections (Table 4.3). The majority of these were between SNPs within the same marker, e.g. the two SNPs within prolactin. Remarkably however, not all SNPs within the same locus showed statistically significant linkage: the two SNPs within the melanocortin-4 receptor (M1046) were not statistically significantly linked, nor were all markers within the growth hormone. Most notably, GH2\_74 and GH2\_60 were not significantly linked, despite being only 14bp apart and despite significant linkage of both markers with GH2\_165, 100bp further on in the gene. Closer investigation of the five SNPs within the growth hormone revealed that a minimum of nine distinct haplotypes are present for the growth hormone sequence comprising these five SNPs (Appendix V).

**Table 4.3.** Benjamini-Hochberg corrected *p*-values for significant linkage between candidate markers in the wild-caught (F0) males.

<b>GH1 &amp; GH2_60</b>	0.0029
<b>GH1 &amp; GH2_211</b>	0.0000
<b>GH2_60 &amp; GH2_165</b>	0.0000
<b>GH2_60 &amp; GH2_211</b>	0.0403
<b>GH2_74 &amp; GH2_165</b>	0.0168
<b>GH2_165 &amp; GH2_211</b>	0.0029
<b>M987 &amp; M1046_Dreyer</b>	0.0123
<b>Prolactin1 &amp; 2</b>	0.0000

#### Genetic diversity

Genetic diversity within samples was moderate and lower for candidate markers than microsatellites: ranging from 0.38-0.47 to 0.40-0.49 over all loci excluding and including microsatellite locus Pr39 respectively (Table 4.4). Measured over all loci, no significant differences in  $H_0$  or  $H_s$  were observed

between lines or treatments within generations or over the F5 and F6 combined (nested ANOVA,  $p > 0.152$ ). Over generations, neither  $H_0$  nor  $H_s$  changed significantly

in the control line (linear regression.  $R^2 = 0.017$  and  $0.000$ ,  $F_{1,62} = 1.05$  and  $0.014$  and  $p = 0.309$  and  $0.907$  for  $H_0$  and  $H_S$  respectively). Furthermore, between the F5 and F6 generations, no significant differences were observed between  $H_0$  and  $H_S$  in selection lines (2-sample randomisation test with 10.000 randomisations,  $p = 0.440$ ;  $0.841$  for  $H_0$  and  $0.910$ ;  $0.692$  for  $H_S$  in small- and large selection treatments respectively).

**Table. 4.4.** Observed ( $H_0$ ) and expected ( $H_S$ ) heterozygosity for candidate loci in generations F0, F2, F5 and F6.  $H_0$  and  $H_S$  are shown per line and generation per locus and over all loci. Standard errors over all loci are obtained through jackknifing over loci. Over all loci estimates are provided both including and excluding locus Pr39, since this locus has many more alleles than SNP markers and may thereby disproportionately affect the estimate of the mean.

	F0		F2		F5_S1		F5_S2		F5_C		F5_L1		F5_L2	
	$H_0$	$H_S$	$H_0$	$H_S$	$H_0$	$H_S$	$H_0$	$H_S$	$H_0$	$H_S$	$H_0$	$H_S$	$H_0$	$H_S$
MB_403	0.279	0.409	0.467	0.421	0.244	0.386	0.476	0.477	0.349	0.374	0.133	0.164	0.204	0.217
M1046_Dreyer	0.686	0.466	0.462	0.422	0.356	0.434	0.5	0.483	0.512	0.459	0.273	0.356	0.478	0.471
M1046_2	0.295	0.316	0.251	0.244	0.304	0.29	0.268	0.235	0.143	0.175	0.152	0.246	0.357	0.302
M30_Dreyer	0.477	0.474	0.487	0.5	0.378	0.507	0.548	0.422	0.667	0.499	0.609	0.47	0.583	0.448
GH1	0.395	0.467	0.464	0.437	0.609	0.47	0.595	0.481	0.564	0.503	0.419	0.46	0.51	0.436
GH2_60	0.467	0.481	0.5	0.489	0.662	0.494	0.55	0.46	0.619	0.5	0.5	0.496	0.66	0.504
GH2_74	0.386	0.412	0.273	0.274	0.37	0.366	0.452	0.402	0.262	0.265	0.311	0.265	0.383	0.338
GH2_165	0.5	0.492	0.505	0.476	0.522	0.501	0.561	0.486	0.535	0.478	0.378	0.482	0.646	0.502
GH2_211	0.386	0.431	0.385	0.355	0.4	0.373	0.561	0.462	0.452	0.423	0.341	0.342	0.49	0.393
Myostatin	0.214	0.265	0.276	0.372	0.261	0.391	0.317	0.393	0.381	0.413	0.378	0.406	0.498	0.456
Prolactin_1	0.511	0.489	0.463	0.461	0.489	0.415	0.571	0.463	0.619	0.486	0.467	0.505	0.479	0.505
Prolactin_2	0.457	0.494	0.468	0.501	0.279	0.502	0.548	0.498	0.5	0.492	0.565	0.501	0.408	0.489
TBC_1	0.4	0.505	0.527	0.501	0.568	0.492	0.512	0.506	0.512	0.459	0.674	0.492	0.458	0.498
SF1	0.182	0.204	0.274	0.252	0.205	0.255	0.095	0.092	0.341	0.318	0.391	0.343	0.286	0.276
M987	0.356	0.374	0.489	0.548	0.413	0.356	0.581	0.607	0.721	0.557	0.804	0.59	0.918	0.621
M9	0.244	0.25	0.255	0.255	0.304	0.26	0.372	0.306	0.233	0.208	0.152	0.142	0.102	0.098
Over all	<b>0.387</b>	<b>0.408</b>	<b>0.409</b>	<b>0.407</b>	<b>0.399</b>	<b>0.405</b>	<b>0.469</b>	<b>0.42</b>	<b>0.463</b>	<b>0.413</b>	<b>0.409</b>	<b>0.391</b>	<b>0.463</b>	<b>0.41</b>
St. err. through jackknifing over all loci	<b>0.0305</b>	<b>0.0245</b>	<b>0.0261</b>	<b>0.0256</b>	<b>0.0349</b>	<b>0.0217</b>	<b>0.0348</b>	<b>0.0312</b>	<b>0.0413</b>	<b>0.0285</b>	<b>0.0477</b>	<b>0.033</b>	<b>0.0477</b>	<b>0.0351</b>
Pr39	0.696	0.676	0.768	0.758	0.739	0.726	0.698	0.636	0.814	0.706	0.87	0.732	0.837	0.709
Pr39 A/ Ar	<b>10</b>	<b>9.206</b>	<b>10</b>	<b>8</b>	<b>6</b>	<b>5.696</b>	<b>8</b>	<b>7.722</b>	<b>7</b>	<b>6.677</b>	<b>5</b>	<b>4.998</b>	<b>7</b>	<b>6.53</b>
Over all incl Pr39	<b>0.405</b>	<b>0.424</b>	<b>0.431</b>	<b>0.427</b>	<b>0.419</b>	<b>0.424</b>	<b>0.483</b>	<b>0.433</b>	<b>0.484</b>	<b>0.43</b>	<b>0.436</b>	<b>0.411</b>	<b>0.485</b>	<b>0.427</b>
St. err. over all incl Pr39	<b>0.0339</b>	<b>0.0279</b>	<b>0.0331</b>	<b>0.0317</b>	<b>0.0384</b>	<b>0.0277</b>	<b>0.0353</b>	<b>0.0319</b>	<b>0.0439</b>	<b>0.0319</b>	<b>0.0523</b>	<b>0.0369</b>	<b>0.0499</b>	<b>0.0357</b>

**Table 4.4.** (continued)

	F6_S1		F6_S2		F6_C		F6_L		F6_L2	
	Ho	Hs	Ho	Hs	Ho	Hs	Ho	Hs	Ho	Hs
M9_403	0.449	0.393	0.469	0.492	0.48	0.485	0.28	0.243	0.26	0.258
M1046_Dreyer	0.531	0.443	0.58	0.494	0.44	0.466	0.408	0.429	0.48	0.407
M1046_2	0.42	0.357	0.3	0.285	0.18	0.229	0.327	0.303	0.3	0.285
M30_Dreyer	0.51	0.502	0.327	0.395	0.42	0.398	0.49	0.412	0.5	0.431
GH1	0.42	0.398	0.5	0.505	0.596	0.499	0.49	0.457	0.48	0.453
GH2_60	0.531	0.5	0.34	0.472	0.551	0.491	0.56	0.484	0.52	0.476
GH2_74	0.38	0.358	0.34	0.335	0.271	0.295	0.32	0.271	0.28	0.243
GH2_165	0.5	0.483	0.52	0.453	0.5	0.48	0.62	0.458	0.42	0.481
GH2_211	0.354	0.345	0.49	0.495	0.5	0.415	0.391	0.343	0.44	0.424
Myostatin	0.204	0.375	0.367	0.43	0.3	0.336	0.327	0.374	0.4	0.407
Prolactin_1	0.24	0.409	0.28	0.426	0.469	0.505	0.5	0.471	0.48	0.502
Prolactin_2	0.4	0.503	0.449	0.5	0.468	0.494	0.4	0.454	0.367	0.489
TBC_1	0.56	0.504	0.4	0.505	0.347	0.422	0.5	0.489	0.48	0.453
SF1	0.417	0.332	0.12	0.114	0.24	0.213	0.38	0.358	0.22	0.229
M987	0.46	0.412	0.78	0.645	0.5	0.443	0.68	0.547	0.96	0.565
M9	0.28	0.243	0.06	0.059	0.24	0.213	0.04	0.04	0.16	0.149
<b>Over all</b>	<b>0.416</b>	<b>0.41</b>	<b>0.395</b>	<b>0.413</b>	<b>0.406</b>	<b>0.399</b>	<b>0.42</b>	<b>0.383</b>	<b>0.422</b>	<b>0.391</b>
St. err. through jackknifing over all loci	<b>0.0263</b>	<b>0.0189</b>	<b>0.0432</b>	<b>0.0377</b>	<b>0.0315</b>	<b>0.0268</b>	<b>0.0379</b>	<b>0.0312</b>	<b>0.0453</b>	<b>0.0299</b>
Pr39	0.72	0.751	0.66	0.535	<b>0.7</b>	<b>0.7</b>	0.98	0.71	0.88	0.696
Pr39 A/ Ar	7	6.635	8	6.981	<b>8</b>	<b>7.518</b>	6	5.638	8	7.448
Over all incl Pr39	<b>0.434</b>	<b>0.43</b>	<b>0.411</b>	<b>0.42</b>	<b>0.424</b>	<b>0.417</b>	<b>0.452</b>	<b>0.402</b>	<b>0.449</b>	<b>0.409</b>
St. err. over all incl Pr39	<b>0.0305</b>	<b>0.0268</b>	<b>0.0435</b>	<b>0.0361</b>	<b>0.0342</b>	<b>0.0308</b>	<b>0.0485</b>	<b>0.0351</b>	<b>0.0504</b>	<b>0.0333</b>

### 4.3.3 Hardy-Weinberg disequilibrium and $G_{IS}$

#### Hardy-Weinberg equilibrium

22 out of 204 tests showed significant deviations from HW-equilibrium, only six of which were significant after B-H corrections for multiple tests: myostatin in the F2, M987 in L2\_F5 and F6, and Pr39 in F5 L2 and F6 L1 and L2. Closer investigation revealed that the significant deviation of myostatin was the result of a heterozygote deficiency and deviations of M987 in the L2 line were the result of an excess of heterozygotes of “307” (the allele without an insert present) and the most common allele with an insert: “317”. Significant deviations in L-lines for locus Pr39 are the result of a deficiency in homozygotes for allele “174” and a large excess in heterozygotes of this allele with either allele “178” or “180”.

## $G_{IS}$

To ensure no significant genetic patterning was ignored, e.g. by inflating the type II error rate due to stringent corrections for multiple testing, results from HW-equilibrium tests and  $G_{IS}$  estimates were compared directly to reveal whether deviations in one or the other underlie interesting genetic patterns that may be the result of different selection regimes.

Table 4.5 shows significant  $G_{IS}$  estimates ( $p < 0.05$ ) with different degrees of statistical support. 49 out of 204  $G_{IS}$  estimates were significant (Table 4.5). The expected number of significant estimates (based on a 5% type I error rate) equals ten but out of the initial 49 significant estimates, only 22 did not stand up to corrections for multiple testing, nor were supported by significant HW probability tests. Fourteen tests had significant  $G_{IS}$  values and deviations of HW-equilibrium but not after correcting for multiple testing. However, four of these tests showed putative interesting patterns between treatments, which justifies closer investigation. Prolactin1 revealed significant  $G_{IS}$  values and a heterozygote excess in F6\_S1 and F6\_S2 only, indicating a potential effect of selection for small size on heterozygosity. However, this pattern was not present in the F5. Weak support for significant  $G_{IS}$  values was furthermore observed for Pr39 in F5\_L1 and M987 in F6\_S2 and F6\_L1. As, like in other populations, the heterozygote excess was caused by a deficiency in “174” homozygotes at locus Pr39 in F5\_L1, and a heterozygote excesses for M987, these deficiencies are still of interest, despite weak support.

Significance after B-H correction in one of the two tests remained for a “174” homozygote deficiency in Pr39 in F5\_L2 and a small heterozygote deficiency for prolactin2 in F5\_S1 and for myostatin in F6\_S1. HW deviations and non-zero  $G_{IS}$  values remained significant after B-H corrections in both tests at myostatin in the F2, Pr39 in F6\_L1 and F6\_L2, and at M987 in F5\_L2 and F6\_L2.

**Table.4.5.**  $G_{IS}$  estimates for 17 candidate loci in pre-selection populations F0 and F2 and selected lines and control line in the F5 and F6 generations.  $G_{IS}$  significance was assessed using 9999 permutations. Colour scales indicate the significance of the estimate based on  $G_{IS}$  permutation tests using GenoDive and/or HW probability tests in GenePop and B-H corrections for multiple tests on both.

	<b>MB_403</b>	<b>M1046_ Dreyer</b>	<b>M1046_ 2</b>	<b>M30_Dr</b>	<b>eyer</b>	<b>GH1</b>	<b>GH2_60</b>	<b>GH2_74</b>	<b>GH2_165</b>	<b>GH2_211</b>	<b>MyostatinProLactin1</b>	<b>ProLactin2</b>	<b>TBC_1</b>	<b>SF1</b>	<b>Pr39</b>	<b>MB987</b>	<b>MB9</b>
<b>F0</b>	<b>0.317</b>	-0.365				0.154					<b>0.193</b>		0.207	0.109			
<b>F2</b>											<b>0.259</b>					<b>0.108</b>	
<b>F5_S1</b>	0.367			0.255	-0.299	-0.296	-0.38				0.333	<b>**0.444</b>		0.199			
<b>F5_S2</b>																	
<b>F5_C</b>			0.183	-0.335												-0.294	
<b>F5_L1</b>	0.188	0.234	0.382	-0.296		0.09			0.216				-0.37		-0.188	<b>**0.363</b>	
<b>F5_L2</b>				-0.303			-0.309		-0.287						<b>*0.18</b>	<b>-0.478</b>	-0.043
<b>F6_S1</b>											<b>**0.456</b>	0.413					
<b>F6_S2</b>							0.28					0.342			<b>-0.234</b>	-0.209	
<b>F6_C</b>								0.081									
<b>F6_L1</b>									-0.354						-0.381	-0.243	-0.01
<b>F6_L2</b>												0.249			<b>-0.265</b>	<b>-0.699</b>	-0.077

**Legend:**

significant  $G_{IS}$ , no support from H-W probability test

significant  $G_{IS}$ , supported by H-W probability test, but not significant after B-H correction

**\*significant  $G_{IS}$  but not after B-H corrections , supported by H-W test significant after B-H correction**

**\*\*significant  $G_{IS}$  after B-H corrections , supported by H-W probability test**

significant  $G_{IS}$ , supported by H-W probability test, after B-H corrections for both

significant  $G_{IS}$  after B-H correction for multiple tests, no support from H-W probability test

Four tests retained significant  $G_{IS}$  estimates after B-H corrections but these results were not supported by HW probability tests (Table 4.5). However, these were most likely the result of discrepancies between the different ways of calculating deviations in both tests, as  $G_{IS}$  permutation tests appear sensitive to small deviations from expectations at loci with few alleles, i.e. SNPs. This is exemplified by M9\_403 and myostatin in the F0, both of which had only very small heterozygote deficiencies (five and two individuals out of 42 respectively). For M987, the negative  $G_{IS}$  in F6\_S2 stems from a heterozygote excess of only four individuals out of 50, and the positive  $G_{IS}$  in M987 from an excess of eleven heterozygotes in 192 individuals.

In summary, heterozygosity levels of prolactin1, M987 and Pr39 deviate from random expectations in a pattern corresponding to the selection treatments and show signatures of selection.

#### **4.3.4 Genotypic differentiation**

Out of 340 within-generation pairwise tests for genotypic differentiation, 118 revealed significant differentiation at candidate loci between population pairs, 83 of which were significant after B-H corrections. These included population pairs for all candidate loci except GH2\_74, M1046\_Dreyer, M1046\_2 and myostatin, for which no significant genotypic differentiation was observed between any population pairs.

Table 4.6 shows  $p$ -values for loci for which significant differentiation was observed between at least some population pairs. Full tables can be found in Appendix VI. The majority of loci differ either consistently over both generations or show a stronger signal in the F6 compared to the F5. The two L-lines are remarkably similar and do not differ significantly for any candidate loci. The two S-lines however differ from each other at several loci, in particular in the F6. The table further shows that S2 differs from L-lines and/ or the control as well at these loci, and more differences exist between L-lines and S2 than between L-lines and S1. In particular, S2 differs at locus M30\_Dreyer from all other lines, while the remaining lines are all similar at this locus. No differences between S-lines are observed for prolactin1, but differences between S- and L-lines are observed for this locus. Differentiation at prolactin2 is



observed only in the F6 between S2 and L1, and has weak statistical support ( $p = 0.0320$ ). This locus is therefore not considered a suitable candidate. M9\_403 is significantly different between both L-lines and S2 and the control. The difference between L-lines and S1 at this locus was significant before B-H correction only ( $p = 0.0156-0.0453$ ). Marker Pr39 and M987 are differentiated between most comparisons and, with M9\_403, seem the most likely candidates for selection based on this analysis.

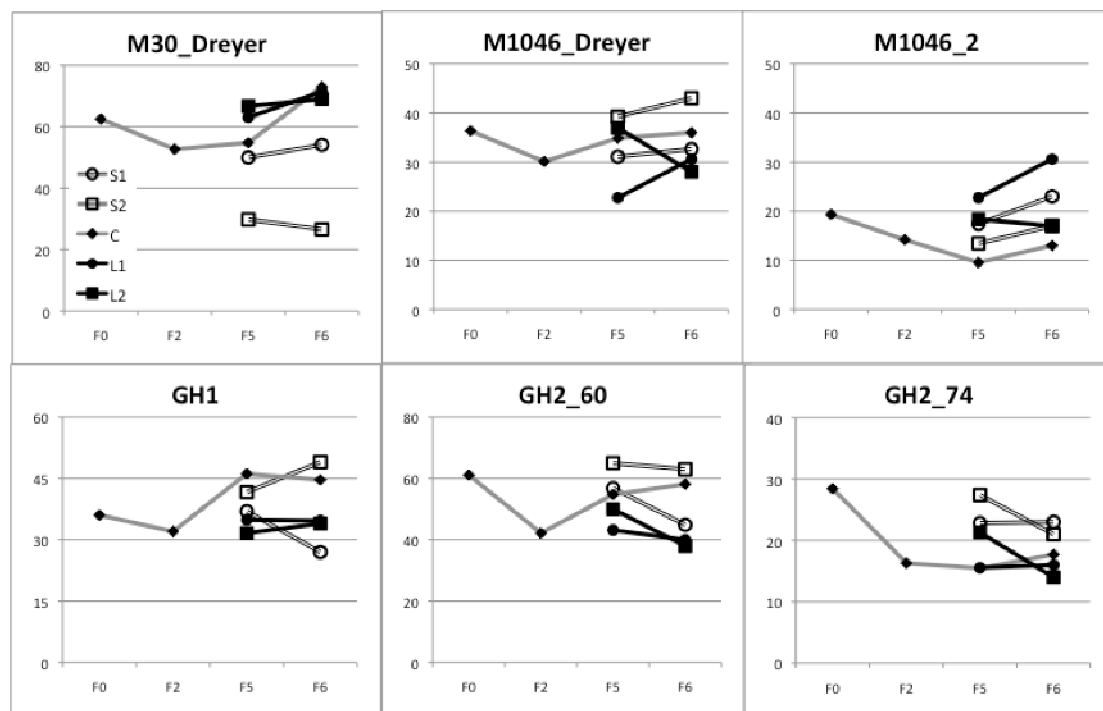
Table 4.6 is based on genotypic differentiation. The allele frequencies underlying these genotypes are summarised in Fig. 4.1, which shows minor allele frequencies

**Table 4.6.** Significant values for pairwise tests of genotypic differentiation, as obtained from a log-likelihood based exact test. Shown are B-H corrected  $p$ -values for each possible population pair within each generation for the different putative candidate loci. Tests between S- and L-lines are shown in grey, within-treatment and between treatments and Control line in white. A significant value for prolactin2 is shown in *Italic*, to highlight the weak support for genotypic differentiation at this locus, which justifies it not being discussed in detail in the text.

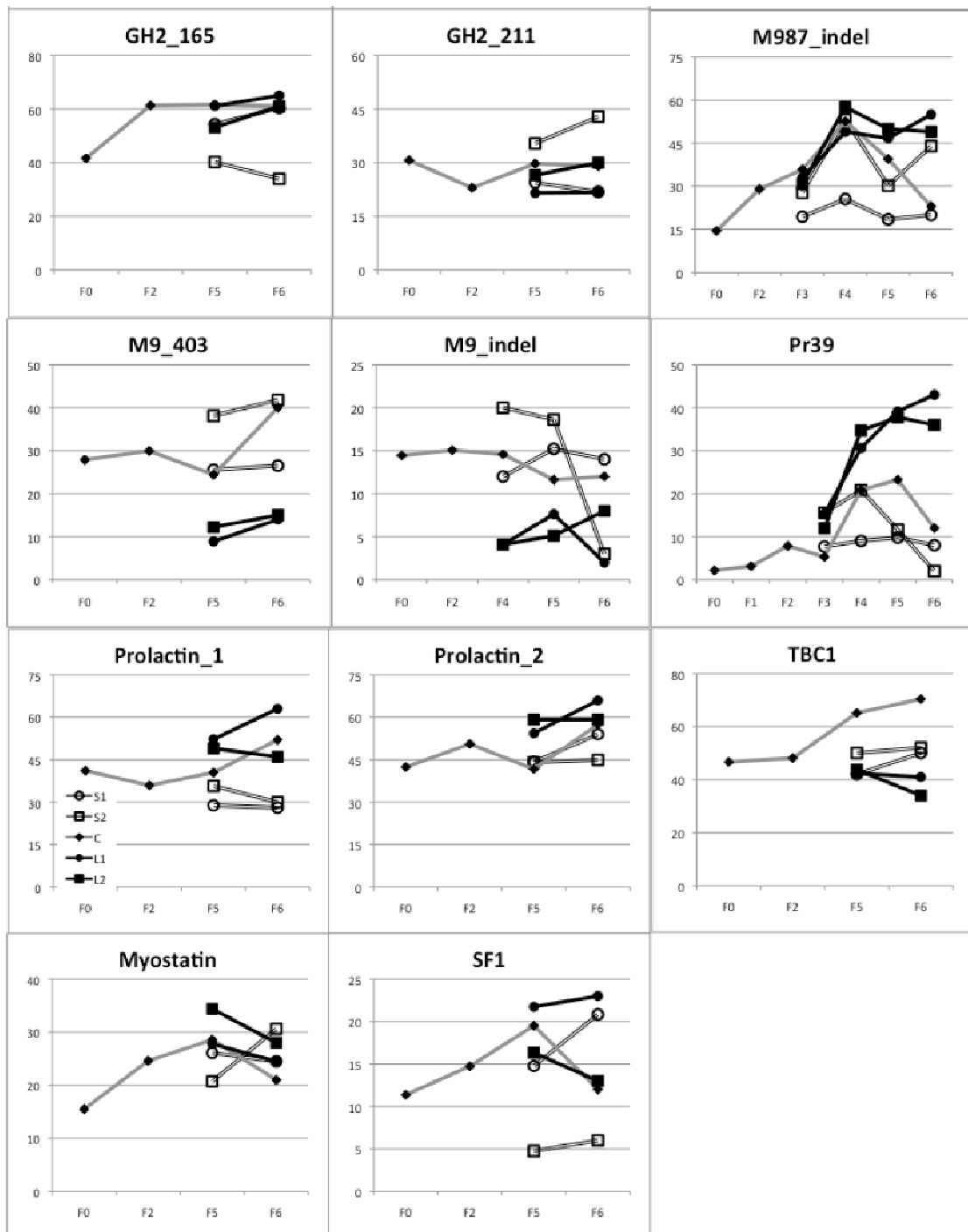
	F5		F6			F5		F6				
	Locus	$p$	Locus	$p$		Locus	$p$	Locus	$p$			
<b>L-L</b>												
<b>S-S</b>			GH1	0.0126	<b>C-S1</b>	TBC_1	0.0084	GH1	0.0321			
			GH2_165	0.0018				Prolactin1	0.0147			
			GH2_211	0.0134				TBC_1	0.0225			
	M30_Dreyer	0.0413	SF1	0.0076				M30_Dreyer	0.0321			
	M987	0.0016	M30_Dreyer	0.0022				Pr39	0.0010			
Pr39	0.0083	M987	0.0000									
		Pr39	0.0000									
<b>S1-L1</b>	Prolactin1	0.0126	Prolactin1	0.0001	<b>C-S2</b>	SF1	0.0175	Prolactin1	0.0276			
	M987	0.0001	M9_indel	0.0109				M30_Dreyer	0.0000			
	Pr39	0.0000	M987	0.0000				M987	0.0000			
<b>S1-L2</b>			Pr39	0.0000	Pr39	0.0002						
	Prolactin1	0.0280	M987	0.0000	<b>C-L1</b>	TBC_1	0.0048	GH2_60	0.0413			
	M987	0.0000	Pr39	0.0000				TBC_1	0.0009			
Pr39	0.0002			M9_indel				0.0326				
<b>S2-L1</b>	GH2_60	0.0206	GH2_60	0.0147				M9_403	0.0007			
	GH2_165	0.0484	GH2_165	0.0000				M987	0.0000			
			GH2_211	0.0109	Pr39	0.0002						
			Prolactin1	0.0004								
			<i>Prolactin2</i>	<i>0.0320</i>								
	SF1	0.0042	SF1	0.0058								
	M9_403	0.0001	M9_403	0.0002	<b>C-L2</b>	TBC_1	0.0245	GH2_60	0.0225			
M30_Dreyer	0.0000	M30_Dreyer	0.0000	TBC_1				0.0000				
M987	0.0022	M987	0.0000	M9_403				0.0015				
Pr39	0.0000	Pr39	0.0000	M987	0.0015							
				Pr39	0.0035							
<b>S2-L2</b>			GH2_60	0.0083								
			GH2_165	0.0022								
	M9_indel	0.0145										
	M9_403	0.0010	M9_403	0.0007								
	M30_Dreyer	0.0000	M30_Dreyer	0.0000								
	M987	0.0058	M987	0.0030								
Pr39	0.0019	Pr39	0.0000									

for all bi-allelic loci, frequencies of the allele without an insert for M987 and of allele “174” for Pr39. The figure supports and clarifies the findings in Table 4.6. The control line shows increasing divergence from selected lines at TBC1 over generations, which differs very little between selected lines. Allele frequencies for S2 differ markedly at GH2\_165 and GH2\_211, M30\_dreyer and SF1 suggesting these markers may be under selection in S2. At GH1 and GH2\_60 however, the differences observed between the S2 and other lines are less consistent and drift seems a more likely explanation for changes in allele frequencies. S1 and S2 are similar for prolactin1 and different from other lines, but the magnitude of this difference does not fall outside the range of frequencies of the control line over generations.

The frequency of the insert in M9\_indel was significant only between S1 and L1 in the F6 (corrected  $p = 0.0109$ ) and between S2 and L2 in the F4 and F5 (corr.  $p < 0.0146$ ). The insert frequency of S2 in the F6 is however the same as for L-lines ( $p = 0.1997 - 1.000$ ).



**Fig. 4.1.** Allele frequencies of minor alleles of candidate SNP markers in generations F0, F2, F5 and F6, and insert frequencies of M9\_indel in generations F0, F2, F4-F6, no insert frequencies for M987 in F0, F2-F6 and allele 174 frequencies for Pr39 in F0-F6. The y-axis show relative allele frequencies on a scale of 0-100. L-lines are depicted by closed squares (L1) and circles (L2), S-lines by open circles (S1) and squares (S2) and the Control by a grey line and closed diamonds.



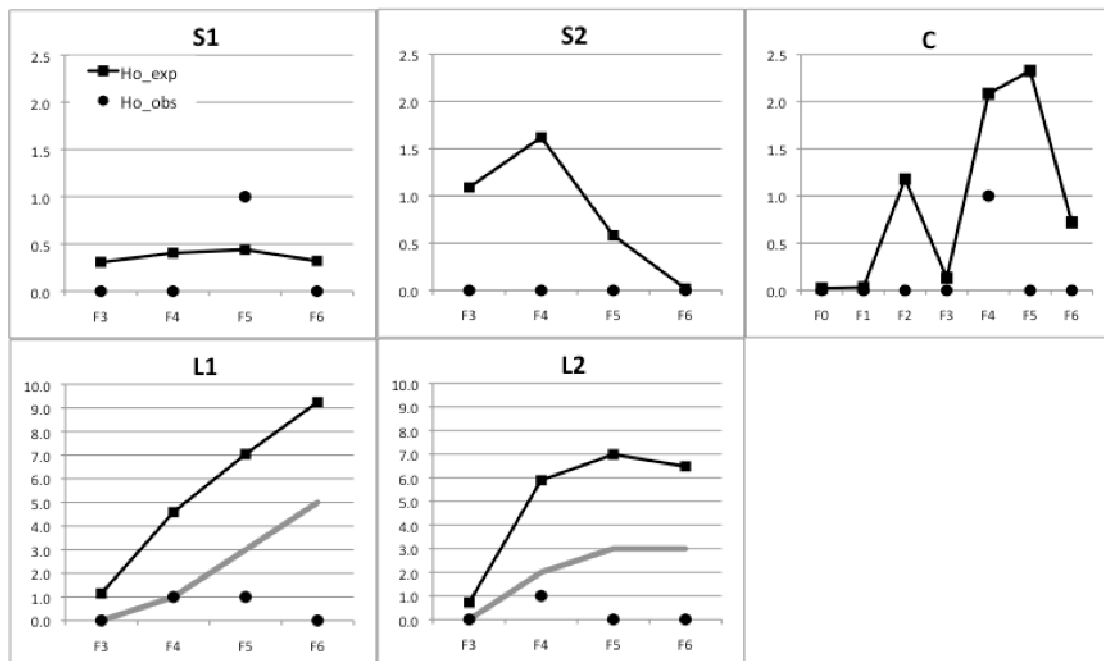
**Fig. 4.1.** (continued)

Whilst differentiation is observed between all lines at locus M987, when simplifying this locus to the presence/ absence of an insert (rather than incorporating all three different insert sizes as for the analysis of Table 4.6), it is only S1 that appears different from other lines, having a markedly lower frequency of the short allele compared to other lines ( $p = 0.000-0.0016$ ) except compared to the control in the F6

( $p = 0.9299$ ). Fig. 4.1 further clarifies that the consistent low frequency in the S1 is not outside the range of frequencies observed in the control line over generations.

No support for significant differentiation was observed for GH2\_74, myostatin, M1046\_Dreyer and M1046\_2. However, from Fig. 4.1 it appears that for GH2\_74, both S-lines have higher minor allele frequencies than other lines. Despite this pattern not being significant this is a noteworthy observation.

For locus Pr39, a strikingly similar increase in frequency of allele 174 is observed for L-lines. Most notably, this increase in frequency is not associated with an increase in homozygotes for this allele (Fig. 4.2). The latter is in particular indicative of selection, because drift is not expected to result in deviations from Hardy-Weinberg equilibrium. This is supported by the microsatellite data in Chapter 3, which revealed no deviations from HW-equilibrium despite drift between lines.



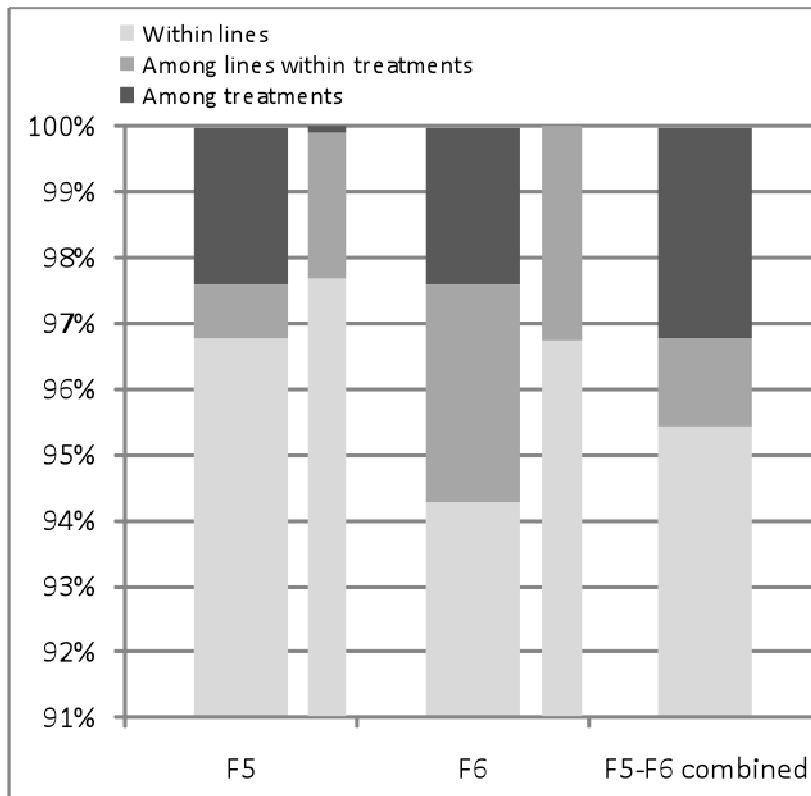
**Fig. 4.2.** Expected (squares) and observed (circles) number of homozygotes for allele 174 for S-lines (S1, S2) Control (C) and L-lines (L1, L2) in generations F3-F6, showing increasing homozygote deficiencies for L-lines but not for S-lines or the Control. The solid grey line in the L1 and L2 graphs represent the cumulative binomial distribution of homozygosity probability for allele 174, based on the allele frequency. Any value of observed homozygosity below the line has a probability less than 0.05. Note the difference in scale on the y-axes between L-lines and the other treatments.

Based on analysis of allele frequencies between lines and treatments, M30\_Dreyer, GH2\_60 and GH2\_165 may be under selection in the S2, M987 in the S1 and TBC1 in the control line. Consistent patterns between treatments are observed for prolactin1, M9\_403 and Pr39, suggesting divergent selection between treatments may be taking place at these loci.

#### **4.3.5 AMOVA**

Analysis of molecular variance (AMOVA) in the F5 and F6 revealed that a significant component of the variation observed in candidate markers was a result of variation between treatments, in contrast to what was observed for microsatellite loci (Fig. 4.3), indicating an effect of the direction of selection on allele frequency distributions at candidate loci. Although the among-treatments component of variation was significant in both generations and identical in size ( $F_{CT} = 0.024$ ), it was less strongly supported in the F6 generation ( $p = 0.0005$  and  $0.036$  respectively). A greater proportion of variation at the within-treatments, among-lines level in the F6 may explain this pattern.

Since not all candidate loci are likely to respond to selection, it is important to reveal which loci are responsible for the observed elevated level of variation between treatments compared to microsatellite loci. However, looking at individual loci, no significant values for  $F_{CT}$  were observed in each generation independently, although  $F_{CT}$  for TBC1 was nearly significant ( $F_{CT} = 0.027$  and  $0.0685$ , and  $p = 0.068$  and  $0.065$  in the F5 and F6 respectively). When statistical power was increased by examining the partitioning of variance in both generations combined, the overall pattern did not differ dramatically (Fig. 4.3) but several loci provide independent statistically significant contributions to the variance observed (Table 4.7). Most notably, the variance components for GH2\_74, prolactin1, prolactin2 and TBC1 are significant at the among-treatments level, but not at the among-lines, within-treatments level. Complete AMOVA tables with variance components and accompanying F statistics for each locus are provided in Appendix III.



**Fig.4.3.** AMOVA of candidate loci in generations F5 and F6, partitioning variance among treatments (black bars) among lines within treatments (dark grey bars) and within lines (light grey bars). Partitioning of variance in microsatellite loci is depicted in the thin bars for comparison.

In summary, AMOVA suggests that selection may operate at GH2\_74, M30\_Dreyer, TBC1, prolactin1, prolactin2, M9\_403 and Pr39, as for these loci significant differentiation between treatments was observed.

#### **4.3.6 Outlier analyses: multiple-marker-based neutrality tests**

##### **Pairwise differentiation**

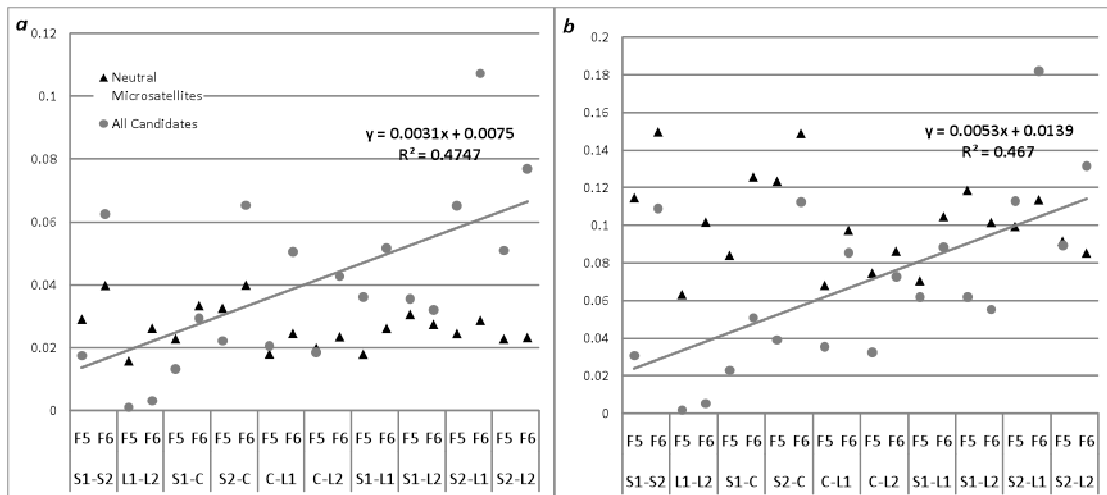
Fig. 4.4 shows pairwise levels of differentiation over all microsatellite and candidate loci between all pairs of population in the F5 and F6 generations. Comparing both graphs, showing  $F_{ST}$  (a) and  $G''_{ST}$  (b) respectively, two observations are immediately obvious: firstly, candidate loci show a strong increase in levels of differentiation from within-treatment to comparisons with the control line and between-treatments

comparisons, which is not observed for neutral microsatellite loci. Secondly, standardisation removes the majority of the difference between microsatellite- and candidate loci, although the gradient over treatments is still only observed for candidates.

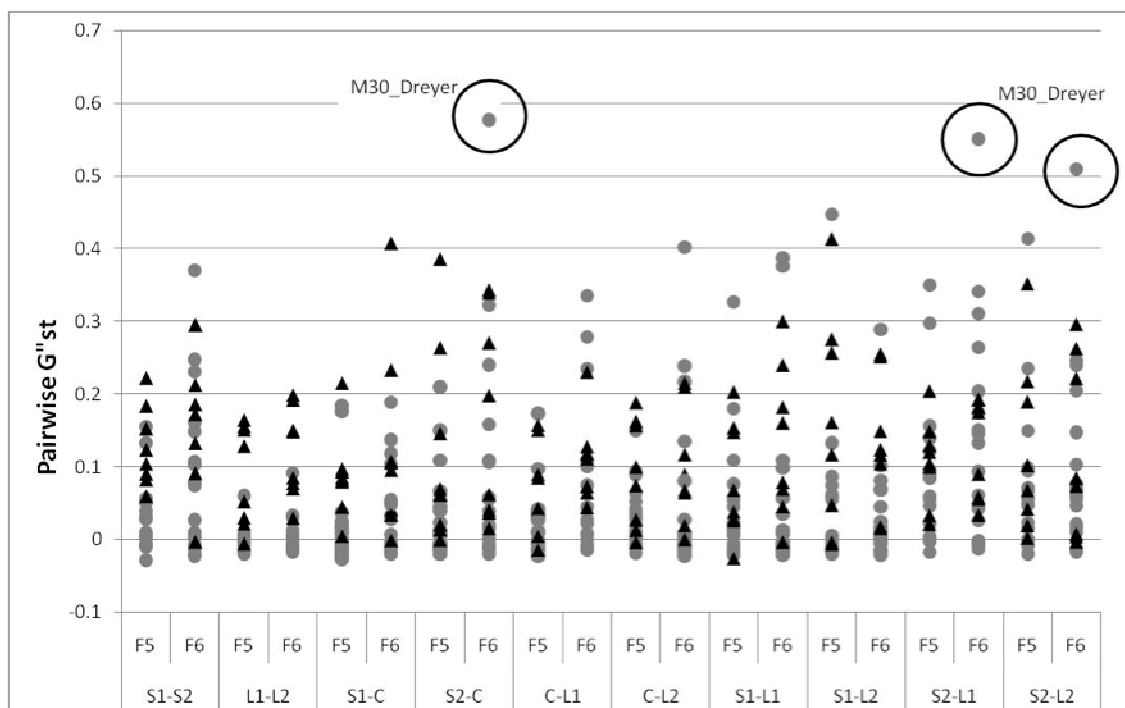
Pairwise  $G''_{ST}$  per locus is provided in Fig. 4.5 and reveals that high  $G''_{ST}$  is observed in particular for M30\_Dreyer between populations S2 and C, L1 and L2 in the F6. No other loci are suggested to be under selection by this analysis. Full tables of both pairwise  $F_{ST}$  and  $G''_{ST}$  are provided in Appendix IV.

**Table 4.7.** Values of  $F_{CT}$ ,  $F_{SC}$ , and  $F_{ST}$  values accompanying Fig. 4.7 with significance levels as ns (not significant), \* ( $p < 0.05$ ), \*\* ( $p < 0.01$ ) and \*\*\* ( $p < 0.001$ ). Shown are values over all loci for the F5, F6 and both generations combined and values per locus for the joint analysis over both generations only. Full AMOVA tables per locus for all three analyses are provided in Appendix III.

	Among treatments (Fct)	Among lines within treatm. (Fsc)	Within lines (Fst)		Among treatments (Fct)	Among lines within treatm. (Fsc)	Within lines (Fst)
F5	0.02412 ***	0.00823 **	0.03216 ***	F6	0.02405 *	0.03389 ***	0.05713 ***
F5 & F6	0.03217 ***	0.01404 *	0.04576 ***				
<b>Per locus for the F5 and F6 combined:</b>							
M9_403	0.07371 *	0.01712 *	0.08957 ***	Myostatin	-0.00059 ns	-0.00125 ns	-0.00184 ns
M1046_Dreye	0.00343 ns	0.00201 ns	0.00543 ns	Prolactin_1	0.05934 **	0.00597 ns	0.06496 ***
M1046_2	0.00466 ns	-0.00391 ns	0.00076 ns	Prolactin_2	0.01726 *	0.00448 ns	0.02166 *
M30_Dreyer	0.08899 *	0.03798 ***	0.12359 ***	TBC_1	0.06024 **	-0.00299 ns	0.05743 ***
GH1	0.00634 ns	0.00634 ns	0.01265 *	SF1	0.00197 ns	0.01915 **	0.02108 **
GH2_60	0.02455 ns	0.00903 *	0.03336 ***	Pr39	0.05519 **	0.02617 ***	0.07991 ***
GH2_74	0.00937 *	-0.00576 ns	0.00367 ns	M987	0.04379 ns	0.04881 *	0.09047 ***
GH2_165	0.01661 ns	0.02061 **	0.03688 ***	M9_Indel	0.01074 ns	0.01653 **	0.02709 ***
GH2_211	-0.00181 ns	0.01352 *	0.01173 *	Over all loci	0.03217 ***	0.01404 *	0.04576 ***



**Fig 4.4.** Pairwise  $F_{ST}$  (a) and  $G'_{ST}$  (b) over all microsatellite loci (black triangles) and over all putative candidate markers for selection (grey circles) between all pairs of populations in generations F5 and F6. The grey solid line shows a linear relationship between differentiation and treatment for candidate loci; from within-treatment comparisons (left) to comparisons with the Control line (middle) and between treatments (right). For microsatellites, such relationship is not visible.



**Fig 4.5.** Pairwise  $G'_{ST}$  values for individual microsatellite loci (black triangles) and putative candidates for selection (grey circles) between all pairs of populations in generations F5 and F6. Putative outliers for locus M30\_Dreyer between S2 and C, L1 and L2 in generation F6 are encircled.



## Fdist

Lositan Selection Workbench was used to detect outlier loci following the fdist method (Beaumont and Nichols 1996). Depending on which combination of populations was compared, different loci emerged as putative outliers and candidates for divergent selection, although considerable overlap in outlying loci was identified.

In the F5, only locus M987 and microsatellite locus G82 were identified as potential candidates for selection in population pairs S1-L2 and C-S2 respectively (Table 4.8). There was no evidence of selection on these loci in the F6. However, when both generations were combined, evidence of selection on both loci was supported.

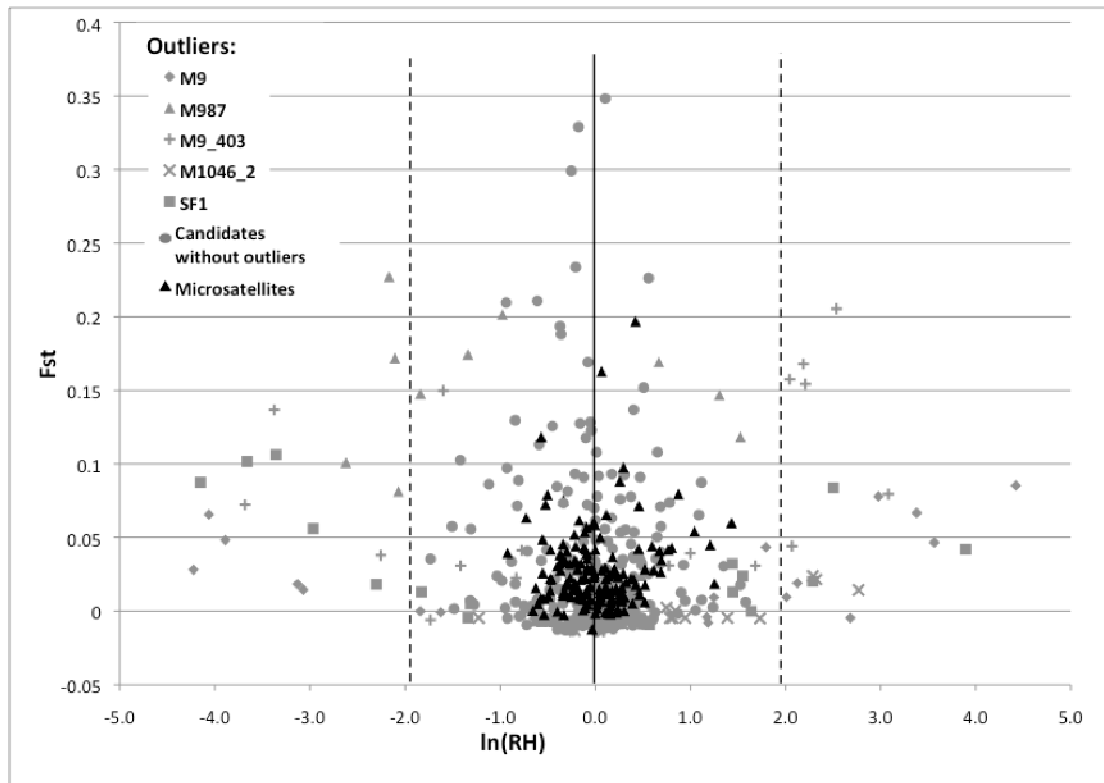
**Table 4.8.** Summary of significant outliers detected by the fdist method using LOSITAN. The table shows expected heterozygosity,  $F_{ST}$  and  $p$  value for all significant loci in the different analyses, with rows showing different pairwise comparisons (between individual populations, treatments and over all) and divided over columns according to generation (F5, F6 or F5 and F6 combined). Grey rows show L-S comparisons.  $P$  values are given as: 1 – the chance that simulated  $F_{ST} < \text{sample } F_{ST}$ . Output figures from all analyses are provided in Appendix VII.

	F5				F6				F5 & F6			
	Locus	He_exp	Fst	p	Locus	He_exp	Fst	p	Locus	He_exp	Fst	p
L-L												
S-S												
S1-L1									Prolactin1	0.5054	0.1105	0.0188
									M987	0.7867	0.0711	0.0234
S1-L2									Pr39	0.5501	0.1316	0.0058
	M987	0.6335	0.2259	0.0065					M987	0.5651	0.1316	0.0203
S2-L1					M30_Dreyer	0.6006	0.3288	0.0217	M30_Dreyer	0.5251	0.1904	0.0120
									Pr39	0.7417	0.1173	0.0164
S2-L2					M30_Dreyer	0.5892	0.2990	0.0149	M30_Dreyer	0.5256	0.1922	0.0011
C-S1									Pr92	0.7581	0.0603	0.0041
C-S2					M30_Dreyer	0.6080	0.3486	0.0107	M30_Dreyer	0.5210	0.1762	0.0087
	G82	0.6099	0.1962	0.0093					G82	0.5806	0.1257	0.0138
C-L1												
C-L2					TBC1	0.5653	0.2267	0.0082	M9_403	0.3611	0.0772	0.0074
									TBC1	0.5127	0.1071	0.0000
									G82	0.5560	0.0613	0.0239
									M987	0.5846	0.0608	0.0101
C-L					TBC1	0.5244	0.1336	0.0181	TBC1	0.5071	0.0752	0.0037
									M9_403	0.3133	0.0745	0.0073
C-S					M30_Dreyer	0.5361	0.1956	0.0201	M30_Dreyer	0.5093	0.1085	0.0064
L-S					M30_Dreyer	0.5157	0.1565	0.0209	M30_Dreyer	0.5047	0.1111	0.0016
					Pr39	0.7587	0.1113	0.0183	Pr39	0.7455	0.0772	0.0030
									M987	0.5946	0.0841	0.0073
All					M30_Dreyer	0.4998	0.1443	0.0077	M30_Dreyer	0.4987	0.0999	0.0027
									Pr39	0.7382	0.0638	0.0095
									M987	0.5808	0.0771	0.0043

In the F6 generation, evidence for selection at locus M30\_Dreyer was obtained for comparisons between S2 and control- and L-lines, but not between S1 and the other populations. Over both S-lines however, a signal for selection at this locus was still identified between S and both control- and L-lines. Support for selection at locus TBC1 and Pr39 was obtained between the control- and L-lines and at locus Pr39 furthermore between S- and L-lines combined. Over all five lines, only locus M30\_Dreyer was identified as a candidate for selection.

In analyses where both generations were combined, the same loci were identified as candidates for divergent selection as in both generations separately. Additionally, the analyses indicated selection at M987 and prolactin1 for S1 and L1 and at Pr39 between several S- and L-line combinations (Table 4.8). Additionally, evidence for selection was identified at microsatellite locus Pr92 between the control line and S1, and at M9\_403 between control line and L2 and both L-lines combined. Between the control and L2, M987 and microsatellite locus G82 were identified as candidates for selection besides TBC1 and M9\_403.

Using the fdist method, M30\_Dreyer thus was detected as an outlier locus comparing S2 to the other lines, and TBC1 only when comparing C and L2. M987 and Pr39 were identified as outlier loci in most F6 and F5 + F6 comparisons. No evidence for selection at any loci was supported for within-treatment comparisons (L-L and S-S comparisons).



**Fig. 4.6.** Plot of  $\ln(RH)$  and  $F_{ST}$  of microsatellite loci (black triangles) and putative candidate genes (grey symbols), showing outliers outside the 95% confidence limits (-1.96 – 1.96, black dotted lines).  $\ln(RH)$  was calculated as the natural logarithm of  $(He_{pop1}/He_{pop2})$  and normalised over all 25 loci per pairwise comparison to mean = 0 and st. deviation = 1.

### Schlötterer’s method

Schlötterer’s method identified several loci as significant outliers. Fig. 4.6 depicts the distribution of  $\ln(RH)$  for all loci between all population pairs and shows the outliers from the distributions, which are summarised in the legend and Table 4.9. It is clear from Fig. 4.6 that microsatellite loci have a much narrower distribution of  $\ln(RH)$  values than candidate loci. Furthermore, the sign of the  $\ln(RH)$  value indicates which population of the pair has the reduced heterozygosity; a negative sign implying greater  $He$  for the population in the right side of the table, and a positive sign greater  $He$  for the population on the left (e.g. a negative  $\ln(RH)$  for pair S1-L1 indicates reduced heterozygosity in S1 compared to L1 for the specific locus).

Table 4.9 shows that the five loci which show outlying values do so in a range of population pairs, not all in concordance with expected differences between selection regimes. For example, locus M9\_indel demonstrates higher than expected  $\ln(RH)$

values between- and within treatments, and no pattern is visible that indicates selection at this locus is the result of different selection regimes in different lines. Similarly, for locus M1046\_2, there is some support for reduced heterozygosity in the control compared to selection lines, but this is only observed between the control and S1 and F5\_L2, therefore providing only little support for selection at this locus. For locus M987, the analysis indicates that S1 has reduced diversity compared to all other lines, but in the F5 only. In the F6, this difference is no longer visible. For locus SF1, the analysis indicates reduced heterozygosity in the S2 when compared to S1, as well as to both L-lines, and the control in the F5. In the F6, S1 has greater diversity than the control line at this locus. For M9\_403 some indication of an effect of treatment on diversity is observed: no support for selection is obtained for within-treatment comparisons or comparisons between the control and S-lines, but L-lines are less heterozygous than both control and S-lines, with exception of S1 in the F6. This is the only locus for which this analysis suggests differentiation to be a result of the selection regime. Full tables of all (normalised)  $\ln(RH)$  values are provided in Appendix VIII.

**Table 4.9.** Summary of loci having  $\ln(RH)$  values falling outside the 95% confidence limits.  $\ln(RH)$  was calculated as the natural logarithm of  $He_{pop1}/He_{pop2}$ , pop1 being the population on the left of the pair, and pop2 the population on the right, and normalised per pairwise comparison to mean = 0 and standard deviation = 1 over all 25 loci. The table shows negative outliers in light grey (indicative of a selective sweep in the population on the left of the pair) and positive outliers in dark grey (selective sweep in the population on the right of the pair). Full tables of (normalised)  $\ln(RH)$  values of all microsatellite and candidate loci are provided in Appendix VIII.

		M9_indel	M987	M9_403	M1046_2	SF1
F5	S1-S2	-0.6462	-2.0699	-0.8335	0.7871	3.8949
F6	S1-S2	3.3790	-1.3399	-0.7738	0.3653	2.5015
F5	L1-L2	2.6813	-0.1820	-1.7344	-1.2260	1.6398
F6	L1-L2	-4.2258	-0.0215	-0.1117	0.2853	1.5491
F5	S1-C	1.1725	-2.6220	0.0889	2.7675	-1.3387
F6	S1-C	0.5197	-0.6372	-1.4174	2.2878	2.2868
F5	S2-C	1.2467	0.2505	0.7727	0.9438	-4.1524
F6	S2-C	-3.8904	1.3043	0.1731	0.8131	-1.8285
F5	L1-C	-1.6258	0.4157	-3.6832	1.7310	0.5000
F6	L1-C	-4.0632	0.6671	-1.5984	0.8407	1.4416
F5	L2-C	-3.1313	0.4887	-2.2562	2.3248	-0.5637
F6	L2-C	-1.8409	1.5254	-3.3758	1.3882	0.5680
F5	S1-L1	2.1235	-2.1078	3.0808	0.4452	-1.3122
F6	S1-L1	4.4216	-0.9780	0.9985	0.1793	-0.4399
F5	S1-L2	3.5638	-2.1697	2.0678	-0.2393	-0.3837
F6	S1-L2	2.0097	-1.8398	1.6833	0.7478	1.4468
F5	S2-L1	1.7933	-0.0346	2.5347	-0.2180	-3.3583
F6	S2-L1	1.1856	0.4777	2.1872	-0.2371	-3.6620
F5	S2-L2	2.9764	-0.1085	2.0441	-0.7142	-2.9660
F6	S2-L2	-3.0703	0.4852	2.2082	0.0405	-2.3017

#### **4.3.7 Association analysis**

Using SNPstats, several associations between male standard length and candidate alleles were found, and for several loci, multiple inheritance models (e.g. co-dominant, recessive) remained significant after corrections for multiple testing. Here, I focus on those models with best support in the whole dataset and the control line and only the most likely models (lowest  $p$ -value and AIC, support in most analyses) are discussed. Complete analysis tables can be found in Appendix IX.

Fig. 4.7 shows the most likely inheritance models for those SNPs and analyses for which a significant association was observed. Strongest support was observed for an association between M987 and standard length and individuals with an insert were significantly smaller than individuals without an insert (corrected  $p < 0.0007$ ). This observation held over all populations (F0, F2, all F5 and all F6), in the F5 and F6 separately, as well as in the control line, indicating that the association is not only a result of divergence in allele frequencies in selection lines, but is also present within (random breeding) populations.

A similarly strong association was found for allele 174 and standard length over all populations, in the F5 and F6 (corr.  $p < 0.00016$ ), 174 being associated with large size. Although this association was not significant for the control line, data for the control is shown here nevertheless (black bars) as it shows the same pattern despite being non-significant ( $p = 0.069$  before B-H correction). As only two homozygotes for allele 174 were observed in the entire dataset, a distinction between dominance or co-dominance of the allele could not be resolved. The latter was also the case for M9\_indel, for which only two homozygotes were present in the data as well, and for which a significant association between the insert and small size was observed overall and in the F5 (corr.  $p = 0.0196$  and  $0.0012$  respectively).

For locus TBC1, no consistent support for a specific mode of inheritance was observed for the association between standard length and SNP genotype. In the F6, only dominance (GG > GT or TT,  $p = 0.0088$ ) was significantly supported after B-H correction. In the control, overdominance of the heterozygote ( $p = 0.0316$ ). Over all

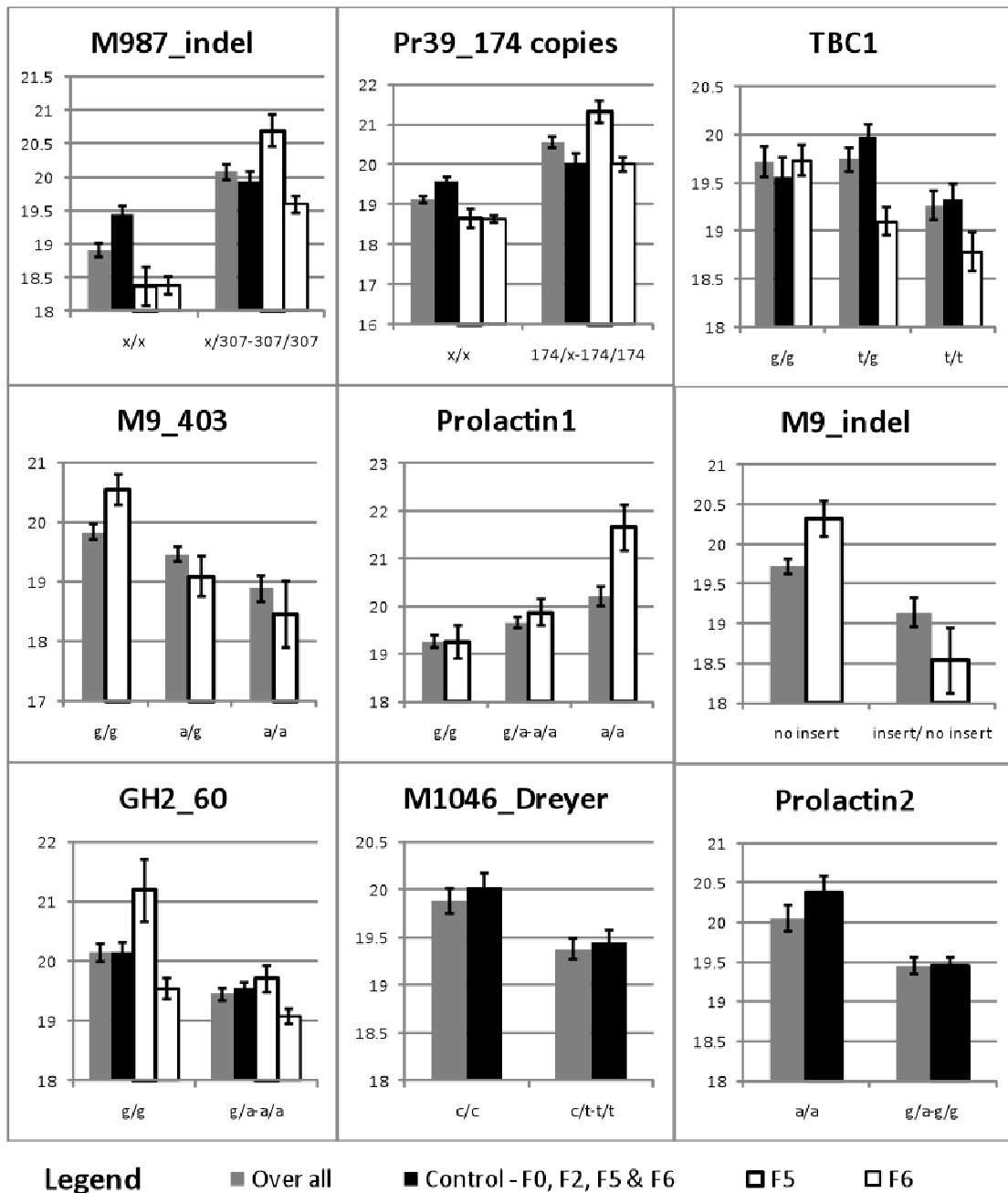
populations, only support for a recessive model (TT smaller than GT and GG) was obtained ( $p = 0.0305$ ).

A log-additive<sup>3</sup> association between M9\_403 and standard length was observed overall ( $p = 0.0036$ ) and in the F5 but in the F5 equal support for dominance of G was observed ( $p = 0.0006$  and AIC = 913.3 and 913.4 respectively). A log-additive model was also supported overall and in the F5 for prolactin1, also with equal support for dominance of one allele in the F5 ( $p = 0.0007$  overall,  $p = 0.0025$  and AIC = 916.7 for both models in the F5).

For M30\_Dreyer, SF1 and GH2\_60, all suggested to be under selection in the S2 in analyses elsewhere (see above), an association was observed over all data and in the F5 for M30\_Dreyer (C recessive overall,  $p = 0.0008$ , co-dominance in the F5,  $p = 0.0001$ ) and overall only for SF1 (overdominance,  $p = 0.0339$ ). For GH2\_60, a significant association was observed over all data and in the control (log-additive,  $p = 0.0007$  and  $0.0390$ ). Equal support for dominance of G was observed in the control ( $p = 0.0390$  and AIC = 1093 and 193.3 respectively). In the F5 dominance of G was also the mode of inheritance with strongest statistical support, ( $p = 0.0320$ ). In the F6, this association was significant but not after B-H corrections ( $p = 0.0430$  before corrections). GH2\_60 may be under selection in the S2 only, as indicated by previous analyses (see 4.3.4), but the significance of an association in the control line independently, suggests that the association of this marker with standard length is more universal and not explained by allelic differentiation between S2 and other lines alone.

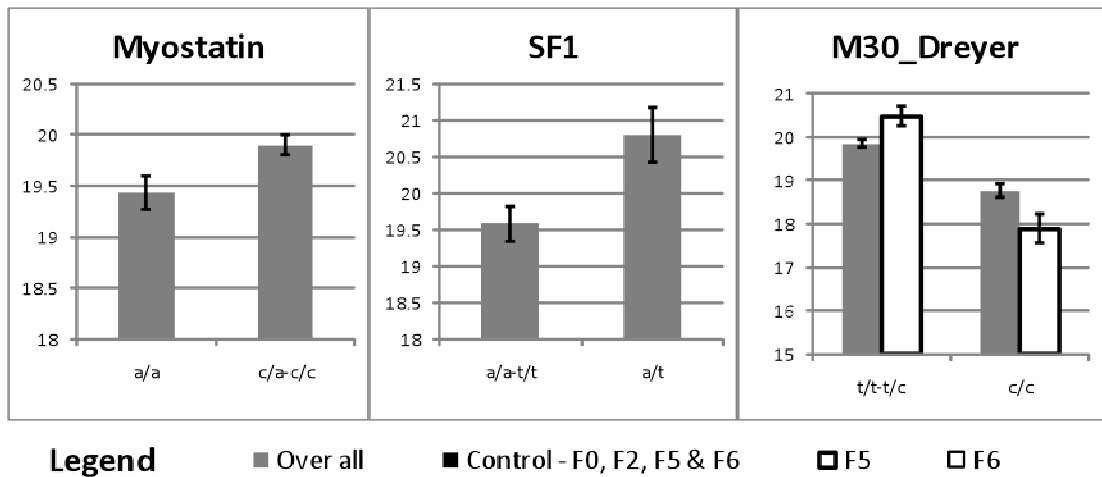
---

<sup>3</sup> Log-additive model: additive effects of each copy of an allele on standard length, showing a logarithmic relationship between copy number and increase/ decrease in trait value.



**Fig. 4.7.** Summary of inheritance models of candidate loci with the strongest statistical support, showing genotype (x-axis) associations with standard length in mm (y-axis) over all data in generations F0, F2, F5 and F6 (grey bars), within the Control line only over generations F0, F2, F5 and F6 (black bars), the F5 (white bars, thick black outline) and the F6 (white bars, thin white outline). Error bars present standard errors of the means. With exception of the Control line for locus Pr39 and the F6 for GH2\_60, only significant associations are shown. All graphs with three bars represent co-dominance of alleles, most graphs with two bars imply dominance/recessivity of one allele, except for locus SF1, which shows overdominance. For locus Pr39 and M9\_indel, too few homozygotes were observed for one allele, which made it impossible to decide on the most likely inheritance model. For locus TBC1, standard length associations are shown for each allele but no conclusion on a most likely inheritance model could be obtained from the analyses. Further details on the graphs can be found in the text. Full association tables are provided in Appendix VIII.





**Fig. 4.7.** (continued).

For M1046\_Dreyer, prolactin2 and myostatin, no convincing support for selection between treatments was observed in any of the other analyses in this chapter, but a significant association between standard length and genotype was nevertheless observed through dominance of one allele for all four loci. For M1046\_Dreyer and prolactin2, the association was observed over all data ( $p = 0.0081$  and  $0.0231$  respectively) and in the control line ( $p = 0.0030$  and  $0.0004$  respectively), suggesting these loci are good candidates for selection but have not been selected for in our lines. For myostatin, weak support was found only over all populations (corrected  $p = 0.015$ ).

## 4.4 CONCLUSIONS & DISCUSSION

### 4.4.1 Evidence for selection at candidate loci

In this chapter, I tried to resolve whether selection over few generations can leave detectable signatures of selection at individual loci. Genetic variation has successfully been characterised at 17 putative candidate markers and detectable rapid evolution over just three generations of selection was revealed in laboratory breeding populations. Several complementary analyses have been used, providing strong support for selection at some markers and ambiguous support for selection at others. Here, the results are briefly summarised.

Table 4.10 summarises the support for selection at different loci for the different analyses in this chapter. From this table, two main observations can be made: first, no unambiguous conclusion on selection for a specific candidate gene can be reached from the analyses, and secondly, there seems to be more variation at candidate loci within the small selected lines than within the large selected lines.

**Table. 4.10.** Summary of analyses for selection detection at putative candidate loci in Chapter 4, showing the magnitude of support for selection at each locus and for each analysis. An “X” indicates strong indications for selection for the particular analysis at that specific locus. Less convincing support, i.e. high *p*-value, no consistent differences between treatments, is indicated by “(x)”. If selection is found for one specific line or population pair only, this is indicated in the Table. Details of all findings and analyses can be found in the text.

	HW diseq/ Gis	Genotypic differentiation	Allele frequency patterning	AMOVA	Pairwise differentiation	Lositan (fdist)	Schlotterer's method (ln(RH))	SNPstats association analysis
Myostatin								
GH1								
GH2_60		X - S2	(x - S2)					X - (S2)
GH2_74				X				
GH2_165		X - S2	X - S2					
GH2_211		(x - S2)	X - S2					
SF1		(x - S2)	X - S2				X - S2	(x - S2)
M30_Dreyer		X - S2	X - S2	X	X - S2	X - S2		(x - S2)
TBC_1		(x - C)	X - C	X		X - C-L2		(x)
M1046_2							(x - C)	
M1046_Dreyer								xx
Prolactin_1	X	X	X	X				(x)
Prolactin_2			(x)	X				xx
M9_indel		(x)	X				(x)	
M9_403		X	X	X			X	(x)
M987	X	X	X			X		X
Pr39	X	X	X	X		X		X
G82							(x)	
Pr92							(x)	

No signatures of selection were observed at myostatin, GH1 and M1046\_Dreyer. Some suggestive evidence was observed for selection at M1046\_2, prolactin2 and GH2\_74. However, this was not observed throughout different analyses and the differences observed at these loci are most likely the result of stochastic change in allelic frequencies between lines and over generations. Visual inspection of the allelic frequencies in Fig. 4.1 clearly shows that the range of allelic differences between treatments at these loci is well within the range observed for the control line over generations. This suggests that drift, rather than selection is responsible for genetic differences between lines at these loci. For TBC1, an obvious pattern for differentiation between the control and selection lines was observed looking at allele frequencies directly in Fig. 4.1. However, statistical support for differentiation at this locus was weak (only in the AMOVA and between L2 and C in LOSITAN) and no conclusive inheritance pattern could be resolved for this locus. Therefore, selection at this locus resulting in differentiation between control and selection lines is not regarded more likely than drift causing the observed pattern of differentiation.

#### **Ambiguous support for selection**

Only few analysis suggested selection at M9\_indel but the allele frequencies at this locus are stable over generations for the control line and, with exception of S2 and the L-lines in the F6, differ between S- and L-lines. The strong decrease in insert frequency in the F6 of S2, may indicate that the magnitude of drift over generations is greater than suspected based on temporal analyses of genetic diversity in 3.3.3 and 4.3.2. However, the stability of frequencies in the four other lines makes this an unlikely explanation. Otherwise, the sudden change in S2 frequencies may reflect that selection for the insert in S-lines has not been sufficient to lead to (near) fixation of the insert in the S2. This rationale stems from the fact that the F5 represent the 10% extreme males from the population, whereas the F6 is a random sample of all male offspring of that generation. Differences between allele frequencies in the F5 and the F6 may thus reflect the different portion of the population being sampled. A third possibility is that divergence between S-lines at this locus represents increased divergence between S-lines and the genetic basis of small size, compared to the genetic basis of large size in L-lines.

At several other loci the observation of increased differentiation between S-lines compared to L-lines is supported. Deviant allele frequencies in the S2 were observed for SF1, M30\_Dreyer, GH2\_165 and GH2\_211 and several analyses identify these loci, in particular M30\_Dreyer, as candidates for selection (Table 4.10). Again, it could be argued that drift has caused divergence of the S2 gene frequencies from other lines. However, except between both GH markers, the analysis did not reveal Linkage Disequilibrium (LD) between any of the markers suggested to be under selection in the S2, which makes drift at all loci causing similar deviations less likely. Furthermore, an association between M30\_Dreyer and standard length has already been demonstrated elsewhere (Tripathi et al. 2009a), making this SNP a highly likely candidate for selection here. Combined, these findings make differential selection between S1 and S2 rather than drift a likely cause for the observed genetic differences at these candidate loci, in particular at M30\_Dreyer.

#### **Polymorphism of haplotypes within the growth hormone**

For the five SNPs in Growth Hormone (GH), a minimum of nine different haplotypes was observed in the F0 generation, which were not in Hardy-Weinberg equilibrium (Appendix V). This is an exceptionally high number of alleles for a coding gene sequence and raises the question whether the sequence as posted on GenBank may contain an intronic fragment, despite it not being classified as such. This indeed proved to be the case (Venkatesh, pers. comm.). Intronic variation can also explain why difficulties were encountered in successful amplification and sequencing of the two additional primer pairs designed for this locus (GH3 and GH4). The presence of an intron is therefore the most likely explanation for the observed variation within the GH sequence, although it cannot be excluded that incomplete replication of fragments during PCR resulted in chimeric products. Polymerase may use such incomplete fragments as templates in subsequent PCR cycles, which can create apparent recombinant alleles (Bradley and Hillis 1997).

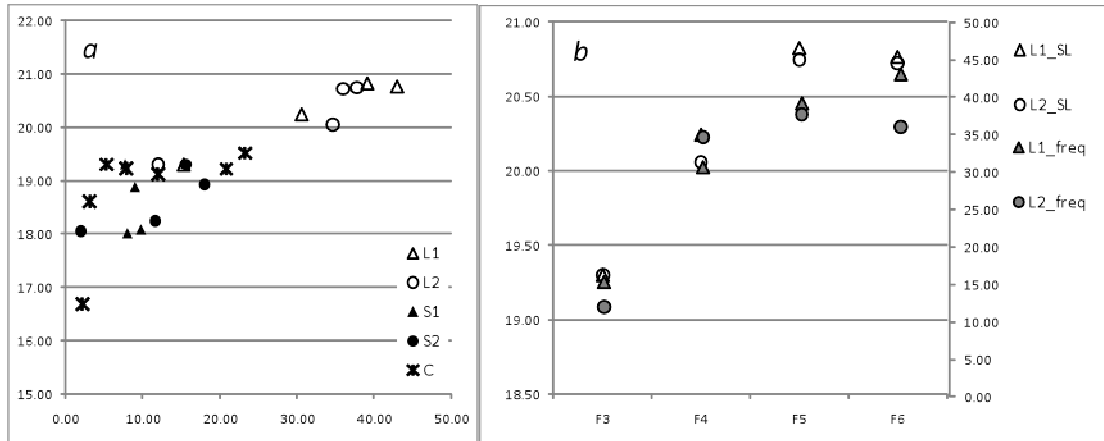
#### **Strong candidates for selection**

Out of 17 putative candidate markers, consistent differences between treatments were observed for M987, Pr39, M9\_403 and prolactin1, which were supported by a

majority of the analyses. For prolactin1, the differences are small, and the association between SNP genotype and standard length at this locus is weak (not supported in the F6 or control line). However, the differences are nevertheless consistent and suggest a role of selection in allelic differences between S-lines and the other treatments.

For locus M9\_403, a significant association between genotype and standard length was only found over all data and within the F5, but work by Dreyer and colleagues (Tripathi et al. 2009a) has previously suggested an association between M9 and standard length, which makes this marker a more likely candidate here. Similarly as at prolactin1, genotypic differences between lines are small but consistent, and visual inspection of allele frequencies clearly revealed similarly divergent frequencies for L-lines compared to S-lines and the control, which were statistically supported.

Locus M987 was identified as a likely locus under selection by the majority of analyses and had significantly higher insert frequencies in S1. HW-deviations in other lines (heterozygote excesses in the S2 and both L-lines) furthermore revealed that the allelic diversity at this locus does not conform to random expectations in other lines. Although Fig. 4.1 reveals that insert frequencies vary substantially over generations within the control line, and drift at this locus thus likely is of considerable magnitude, a strong association between standard length and the absence/ presence of an insert was observed in all association analyses, including the control line independently. This suggests that, despite high stochasticity, this locus, or a tightly linked locus, may be involved in the genetic basis of standard length and selection is likely at least partially responsible for the allelic divergence between S1 and other lines. Furthermore, Tripathi et al. (2009a) also found an association between a SNP in M987 and standard length, providing further support for the relevance of the association found here. The observation of S1 and S2 differing in genotype frequencies at this locus, adds further support to the hypothesis of a different genetic basis to small size in the S1 compared to the S2.



**Fig. 4.8a.** Relationship between 174 allele frequency of locus Pr39 (x-axis) and mean male standard length in mm (y-axis) for L-lines (open symbols), S-lines (closed symbols) and the Control (black asterices), including generations F0-F6.

**b:** Correlation between mean male standard length in mm (left y-axis, open circle and triangle) and allele 174 frequency of Pr39 (right y-axis, closed circle and triangle) in L1 (triangles) and L2 (circles) over generations F3-F6 ( $r = 0.953$ ;  $p < 0.001$ ).

The most convincing support for allele frequencies at individual loci responding to selection on standard length was found for microsatellite locus Pr39. A strong association of the presence of allele “174” with standard length was found, and the frequency of this allele increased in parallel with standard length in L-lines (Fig. 4.8).

Finally, although no evidence for selection at this locus was observed in the analyses, the association between M1046\_Dreyer and male standard length, as observed by Tripathi et al. (2009a), has been confirmed.

#### **4.4.2 Congruence and discrepancies between analyses**

The majority of analyses suggested the same loci as most likely candidates for selection: M987, Pr39, M30\_Dreyer and, although the divergence between lines is less pronounced, M9\_403 and prolactin1. However, in contrast to the other analyses performed, tests for pairwise differentiation and Schlötterer’s method did not put most of these loci forward as likely candidates for selection and only few outlier loci were suggested from these analyses (only M30\_Dreyer by pairwise tests of differentiation and SF1 and M9\_403 by Schlötterer’s method). In fact, not one locus was identified as an outlier by all three outlier methods and only M30\_Dreyer by more than one. What could explain these discrepancies between analyses?

The major difference between the different outlier detection methods is the statistics used to determine outliers. *Fdist* uses a simulated distribution of  $F_{ST}$  values to determine a neutral distribution of  $F_{ST}$ , from which outliers can be identified. For the pairwise tests of differentiation, the neutral distribution is obtained from a comparison between neutral microsatellites and putative candidate genes directly. Thereby, greater stochasticity in defining “neutral” differentiation levels between population pairs is introduced, which can reduce the power to detect outlier loci. The use of  $G''_{ST}$  rather than  $F_{ST}$  as a measure of differentiation, furthermore increased differentiation for microsatellite loci disproportionately compared to SNPs, thereby reducing differences between neutral and putative candidate loci.  $G''_{ST}$  differs from  $F_{ST}$ , in that it corrects for variable levels of heterozygosity affecting maximum differentiation (Meirmans and Hedrick 2011). Using  $F_{ST}$ , SNP markers are expected to have higher levels of differentiation than microsatellites, as a result of lower average heterozygosity levels than found on average at microsatellite loci.  $G''_{ST}$  corrects for this bias by normalising loci over the maximum attainable differentiation level, based on heterozygosity, and is therefore more suitable for comparing differentiation between different classes of loci (i.e. microsatellites and SNPs).

LOSITAN is based on  $F_{ST}$  and does not correct for a potential bias in differentiation due to variable levels of heterozygosity and to my knowledge,  $G''_{ST}$ -based outlier methods following the *fdist* method do not yet exist. A second potential factor biasing outlier detection in LOSITAN is caused by the different mutation models likely to underlie microsatellites and SNP markers. An infinite alleles model (IAM) was used for our analyses, but a (partial) stepwise mutation model (SMM) may be more appropriate for microsatellite loci. This can lead to underestimation of differentiation levels at microsatellite loci (Excoffier et al. 2009) and thereby may have inflated differences between putative candidates and microsatellites. However, despite these limitations, microsatellite locus Pr39 was identified as a candidate locus by *fdist*, suggesting that the use of  $F_{ST}$  rather than  $G''_{ST}$  and an IAM have not majorly biased differentiation measurements between microsatellites and SNP candidate markers and not seriously limited the potential of the *fdist* method to detect outlier loci among the diverse set of markers here used.

Schlötterer's method is based on the expected normal distribution of the natural logarithm of differences in heterozygosity between two populations. Thereby, it differs from the other outlier methods in that it does not use F-statistics to detect differentiation between populations, but is only sensitive to reductions in heterozygosity in one population compared to another (Schlötterer 2002). In particular for Pr39, for which only a single allele is likely to be under selection and reveals an association with standard length, heterozygosity did not differ between populations. Consequentially, selection for this allele remains easily undetected by an analysis based on selective sweeps. When the analysis is done with Pr39 as a bi-allelic locus (so looking at the presence/ absence of allele 174 only), this locus is identified as an outlier in pairwise comparisons between all treatments (Fig. d in Appendix VIII).

Such discrepancies between analyses have been reported elsewhere too. Beaumont & Nichols (1996) and Vitalis et al. (2001) used the same set of data to illustrate the use of different outlier methods (fdist and DetSel respectively) and identified different outlier loci. Vitalis et al. (2001) ascribed this to differences in sensitivity to the inclusion of isolated populations and variation in population size, to which fdist appeared to be more sensitive. However, as populations here are all of the same size and isolated from each other, this is not a concern in our dataset. Martínez et al. (2011) also observed discrepancies between candidate loci retrieved by fdist and  $\ln(RH)$  and attributed these differences to the non-linear relationship between variation in allele frequencies and heterozygosity. The authors furthermore emphasised that, although congruence between different methods provides greater support for true outliers and minimises the likelihood of false positives, due to the different underlying statistics, true outliers can be obtained despite contrasting results in different analyses.

### **Variation between generations**

An interesting observation when comparing results between analyses, is that not all analyses identified the same loci as most likely loci under selection in the same generations. Tests for pairwise differentiation, LOSITAN, HW-equilibrium and,



although the difference is less pronounced, tests for genotypic differentiation all identified higher numbers of loci as putatively under selection in the F6 than in the F5. Using Schlötterer's, the difference between generations was not pronounced but in SNPstats, most significant associations were found in the F5 rather than the F6.

Greater differentiation between the F6 populations compared to the F5 may be expected, since another generation of breeding separates the lines. However, the F6 samples are a random subset of each population, whereas the F5 represent the smallest/ largest 10% of males in the selected lines. The fact that still a greater signal for selection is observed in the F6 may be regarded as an indication of the efficacy of the selection regime in causing genetic change at candidate loci throughout the population, and not just in the extreme percentiles.

In contrast to the other analyses performed, SNPstats identified most significant associations between genotype and standard length in the F5 rather than in the F6. However, patterns between genotypes and standard length were the same in the F6 as in other generations despite being non-significant. The observation of such a dilution of the association in the random samples of the F6 generation, compared to the extreme samples of the F5, most likely stems from the fact that the phenotypes and genotypes in the F5 represent the 10% extremes from each selection line, whilst both are a random sample of populations in the F6. Consequentially, phenotypic distributions are less extreme in the F6 than in the F5 and associations will be weaker.

In the AMOVA, more variation between lines within treatments was observed in the F6 compared to the F5, which was a result of increasing variation between S1 and S2. Increasing genetic variation between S1 and S2, as opposed to L1 and L2, supports the phenotypic results obtained in 2.4.3, which revealed variation in maturation age and size between S-lines but not between L-lines.

### **The power to detect selection**

For all three outlier methods used here, by definition 5% of loci are expected to be outliers if the desired significance level is  $p = 0.05$ . Pairwise differentiation resulted in only one obvious outlier (M30\_Dreyer), which is below the 5% margin and therefore this locus may not be a true outlier but due to random chance. Schlötterer's method revealed five out of 25 loci (20%) to be candidates for selection in at least one population pair and LOSITAN provided eight out of 25 loci (32%), including two microsatellite loci, for which no other analysis suggested selection to be the cause of differentiation (G82 and Pr92). Although the proportion of outliers in both tests is higher than can be expected at random, the identification of the two microsatellites as outliers in a small fraction of the comparisons (C-S2 all for both and C-S2\_F5 as well for G82) suggests that false positives are a real concern for the outlier analyses and emphasises the importance of complementary analyses in detecting selection: the lack of support for selection at G82 or Pr92 in any of the other analyses strengthens the conviction that these two markers represent false positives rather than loci under selection.

A complication with the analysis performed here is that, with exception of the AMOVA, none of the analyses allows for implementation of the nested design of replicate lines within treatments. Ignoring such a hierarchy may result in an inflated type II error rate. Several other aspects of our approach may have affected the power to detect selection in our experiments. First of all, considerable portions of the DNA sequences used to mine SNPs were unreadable, minimising the gene regions in which candidate SNPs could be identified. Secondly, the number of putative candidate markers (17) was greater than the number of putative neutral loci (8), which results in little confidence in determining neutral genetic distributions to compare candidates to. Numerous outlier studies have been previously conducted by others employing tens to hundreds of markers (Nielsen et al. 2009b; Martínez et al. 2011; Russello et al. 2011) consequentially having greater scope and power to detect outliers. Studies using numbers of loci comparable to our own often compensate for this with greater sample sizes (Larsson et al. 2007). Although the sample sizes of 50 individuals used here proved sufficient to prevent major drift and

inbreeding in our lines, it may well limit the power to detect small genetic change over few generations and at low levels of differentiation (Ryman et al. 2006; Excoffier et al. 2009). Furthermore, SNPs have lower power to detect population differentiation than multi-allelic (microsatellite) loci (Xiong and Jin 1999; Ryman et al. 2006). On the other hand, the use of candidate markers for which prior knowledge supports a role in growth and size (either from linkage mapping or known gene function), provides convincing support for associations found here (i.e. M30\_Dreyer, M9\_403, M987, prolactin1).

### **Selection or drift?**

Prior information on the candidate loci used in this study also decreases the likelihood of drift explaining the patterning between selection lines as opposed to selection. However, it is difficult to exclude a role for drift in the observed changes. Comparing the change in allele frequencies of the control to those at selection lines (Fig. 4.1) reveals that, at least for M987 and prolactin1, the magnitude of drift may be considerable. For M30\_Dreyer and Pr39 however, the magnitude of allelic change between treatments is far greater than that observed between generations in the control, supporting the conclusion of selection rather than drift inducing genetic change at these candidate genes.

Further support for a lack of drift affecting candidate gene distributions, is the lack of support for significant trends in heterozygosity over generations within lines or differences between control- and selected lines. This is also visualised by Schlötterer's method, which found little support for selective sweeps occurring, thereby indirectly supporting the conclusion that captive breeding and selection did not cause genetic erosion in breeding lines (Chapter 3).

Perhaps the most convincing support for the effect of selection on genetic variance at candidate loci comes from the AMOVA. Not only was there a significant value for  $F_{CT}$  (variation among treatments) at candidate loci, no significant  $F_{CT}$  was observed for neutral microsatellites (see 3.3.4), illustrating the different genetic processes, i.e. drift and selection, underlying differentiation at neutral and candidate loci. However,

$F_{SC}$  (variation among lines within treatments) was also significant for candidate loci, including Pr39, M987, M30\_Dreyer and M9\_403, but not prolactin1, signifying that drift too is likely to play a considerable role in candidate gene differentiation between lines.

#### **4.4.3 The genetic basis of male body size: differences between small and large size**

The data presented here reveals that the two L-lines are strikingly genetically similar when compared to S-lines, between which differentiation is observed in the F6 in particular. This observation corresponds with the observed differences in size and age of maturation between S1 and S2 described in Chapter 2 (see 2.4.3), which indicated, although not statistically significant, that S2 individuals matured and grew to larger sizes than the S1.

As previously hypothesised in Chapter 2.5.3, greater diversity in the genetic basis of small body size compared to large body size can be expected due to effects of inbreeding and the broader genetic basis of small size compared to large size. Based on the phenotypic data from our experiments, I predicted greater heterozygosity underlying large body size and a higher frequency of recessive homozygotes under small fish.

Over all candidate genes (4.3.2) or at neutral microsatellites (3.3.3), no differences in levels of heterozygosity were observed either over all candidate genes (4.3.2) or at neutral microsatellites (3.3.3). However, looking at individual candidate loci and using the association analysis of putative candidate genes as performed in 4.3.7, some support for heterozygosity being associated with large size is observed. Out of nine associations where standard length was significantly different between homozygotes of one allele and both heterozygotes and homozygotes of the other allele (Fig. 4.7), five revealed an association of small body size with homozygosity (M987, Pr39, myostatin, SF1, M30\_Dreyer) and four revealed homozygotes to be the larger genotypes (M9\_indel, GH2\_60, M1046\_Dreyer, prolactin2). Although overall these findings do not support the hypothesis, the prediction does hold for the strongest candidates for selection; homozygotes for one allele were observed to be

smaller than heterozygotes and homozygotes of the other allele at Pr39, M987, and M30\_Dreyer (Fig. 4.7). For prolactin1 and M9\_403, the other two strong candidates for selection, small body size is associated with homozygotes, but a co-dominant model of inheritance obtained the strongest statistical support. Over these five loci, no significant differences in  $H_0$  or  $H_S$  were observed between S- and L-lines ( $p > 0.181$ ). However, this is mainly an effect of M9\_403. Over the three dominantly inheriting loci and prolactin1, a significant difference in observed heterozygosity ( $H_0$ ) is found between S- and L-lines (independent sample t-test over the F5 and F6 combined;  $T_{(2),30} = -2.554$  and  $T_{(2),22} = 0.842$ ,  $p = 0.016$  and  $0.01$  for tests with and without prolactin1 respectively). This indicates reduced heterozygosity at the most likely candidate loci for selection in S-lines compared to L-lines. No differences in expected heterozygosity ( $H_S$ ) were observed ( $p > 0.290$ ). These data thus indicate that part of the observed variation in standard length between treatments may be associated with differences in heterozygosity.

However, no differences in  $H_0$  or  $H_S$  were observed between S1 and S2 at these loci ( $p > 0.316$ ). Conversely, at M30\_Dreyer, GH2\_60 and SF1; the loci for which support for selection at S2 only was observed (Table 4.10), a significant difference in  $H_0$  was found in the F6, indicating greater heterozygosity in the S1 – the smaller fish – than the S2 ( $T_{(2),4} = 2.817$ ,  $p = 0.048$ ). At the individual loci, homozygosity was associated with large size at GH2\_60, but with small size at M30\_Dreyer (Fig. 4.7). Large size at locus SF1 was associated with the heterozygous genotype over all data, further abating a role of this locus in explaining the differences between S1 and S2. In conclusion, these data suggest that, as can be expected for a quantitative trait, a considerable source of genetic variation underlying variation in standard length between selection lines is unaccounted for in our analyses.

#### **4.4.4 The genetic basis to body size: Y-linkage of candidates**

The phenotypic response to selection, as described in Chapter 2, indicated strong y-linkage of the genes involved in the selection response, as no response was observed in females and heritabilities were high. The genetic data here presented provide further support for this conclusion and suggest Y-linkage of candidate genes through two main observations.

Firstly, it was already known from the work by Tripathi et al. (2009a), that M9, M30 and M987, some of the strongest candidates for selection in our experiments, are located on the sex chromosome in guppies. In addition, significant LD between M1046 and M987 in our study further suggests that M1046, the melanocortin-4 receptor gene, may also be located on this chromosome. This is further supported by knowledge on the location of this gene on the sex chromosome of *Xiphophorus spp.* (Lampert et al. 2010).

Contrasting to sex chromosomes in mammals, those in many fish species, including the guppy, are in early stages of differentiation and can be regarded as pseudo-autosomes: sex-linkage of alleles at marker loci may be observed at loci closely linked to the sex determining locus (SDL) but recombination between the x- and y-chromosome is possible throughout the majority of the pseudo-autosome. Therefore, heterozygosity and Y-linkage – two incompatible terms for mammal species – can be observed in fish (Volff and Schartl 2001; Tripathi et al. 2009b). The guppy SDL specifically has been mapped to the far distal end of the y-chromosome (Tripathi et al. 2009a, 2009b). M9, M30 and M987 can be found at the top end of the chromosome, separated from the SDL by a large repeat-rich region. Tripathi et al. (2009a) observed sex-specific segregation of alleles at these markers, although no allele was exclusively associated with the Y-chromosome. This observation is supported by our findings. For M987, for which an association between standard length and the presence of an insert has been demonstrated here, heterozygotes and homozygotes for both the presence and absence of an insert were observed throughout the dataset. Contrastingly, homozygotes for the insert of M9\_indel, which is located not far (0.9 cM) from M987, were observed in only 8 out of 960

genotypes, providing stronger support for sex linkage of a candidate gene for male standard length located in the near vicinity of these markers.

Y-linkage may also explain the results obtained for Pr39, for which an association with only one specific allele, “174” and standard length was observed, and for which L-lines, where the allele was most frequent, showed a strong heterozygote excess. Homozygotes for 174 were only observed in five individuals out of the entire dataset of 1236 male fish. This suggests that a closely linked gene affecting male size may have a genotype causing large size linked to the almost exclusively Y-linked 174 allele specifically. Although this rationale can explain the heterozygosity excess for 174 and the fact that only a single allele shows the association with standard length, it remains speculative as no female genetic samples are available and therefore Y-linkage of markers could not be tested. A literature search for usage of Pr39 in QTL mapping studies (e.g. Shen et al. 2007b; Tripathi et al. 2009a) or population genetic studies reporting on allele-specific deviations from HW-equilibrium (van Oosterhout et al. 2006a; Barson et al. 2009; Suk & Neff 2009) remained futile.

#### **4.4.5 Putative gene function of major candidate genes**

Out of the five major candidate genes for selection; prolactin, M9, M30, M987 and Pr39, three genes have been assigned a putative function. Besides microsatellite marker Pr39, no function could be assigned to M987, which did not yield any significant hits in a blast search. The presence of insertions of 9, 10 and 11 bp within the sequence of this marker strongly suggests this is a non-coding sequence. Prolactin consists out of an mRNA sequence highly similar to prolactin mRNA of *Tilapia nilotica* (blast E-value:  $1e^{-52}$ ). Thirteen percent of the mRNA sequence M9 resembles rab interacting lysosomal-like protein-1 (RILPL1) in chimpanzee (*Pan troglodytes*, E-value:  $9e^{-47}$ ). In humans, RILPL are neuroprotective proteins, inhibiting the apoptotic function of glyceraldehydes 3-phosphate dehydrogenase (GAPDH) and thereby affecting cell turnover rate (neXtprot database. Available at: <http://www.nextprot.org/db>). In rainbow trout (*O. mykiss*), RILPL is differentially expressed between starved and unstarved fish (Salem et al. 2007). M30 consists out of the mRNA sequence of fructose-1,6-bisphosphatase 1 (FBPase) and closely

resembles FBPAse in, amongst others *Salmo salar*, *Oreochromis niloticus* (E-value: 0.0) and *Danio rerio* (E-value:  $3e^{-155}$ ). FBPAse is an enzyme involved in glucose synthesis and a variety of metabolic pathways affecting energy generation and development (e.g. Douglas et al. 2008). The open reading frames (ORFs) for these markers are unidentified and therefore it is unknown whether the SNPs investigated represent synonymous or non-synonymous mutations.

#### **4.4.6 Overall conclusions**

The data here presented provide evidence for rapid evolutionary change at specific candidate loci, induced by strong directional selection on male standard length. To our knowledge, this study represents one of the first cases where detectable selection at individual loci has been demonstrated in artificial selection system over few generations. I only know of one other example of an artificial selection study (Fidler et al. 2007) where selection at a candidate locus has been revealed over such a short timeframe. However, Fidler et al. (2007) did not use a control line in their selection experiments nor compared candidate gene frequencies with supposedly neutral, stochastic, genetic change between selection lines and generations. This emphasises the unique position of our work, in which the effects of selection have been quantified both at a phenotypic and a genetic level and have been compared to neutral variation.

Out of seventeen putative candidate loci, five revealed substantial support to be responding to selection on standard length in our experiments (Pr39, prolactin1, M9\_403, M30\_Dreyer and M987). For the latter three of these markers, an association with standard length was already established in previous work (Tripathi et al. 2009a) and at M30\_Dreyer specifically, this association was established for the same polymorphism. A previously observed association between the M1046\_Dreyer and standard length could also be confirmed with our data, but selection at this SNP did not occur.

The analyses performed in our study highlight the importance of using a variety of methods in the search for loci under selection: selection may act on heterozygosity,



allele frequencies, HW-equilibrium or a combination of these and our analyses have revealed that different methods are sensitive to different levels of change in these parameters. Especially under the low levels of change that can be expected after only few generations of selection, discriminating between drift and selection can be challenging and is aided by using complementary analyses, replicate populations for different treatments and careful selection of candidate loci, as the present study testifies.

## **5. GENERAL DISCUSSION**

## **5.1 SUMMARY OF MAIN RESULTS AND NOVELTIES OF THE RESEARCH**

The present study represents the first case of size-selective harvesting on selection lines established from wild-caught fish, in which the phenotypic response is compared directly to genetic change at both neutral genetic markers and genes under selection. Our findings revealed that size-selective harvesting can cause significant genetic change and phenotypic shifts in life history traits of captive populations of Trinidadian guppies. Over only three generations of selection, standard length changed by -6.5% and 7.5% in lines selected for small and large body size respectively (Chapter 2). Allele frequencies at the major candidate locus changed by as much as 41% over generations (Pr39, from 0.022 to 0.43 in L-lines) whilst the range of allele frequencies in the random breeding control was only 0.022-0.21 (Chapter 4).

We revealed considerable additive genetic variation for male standard length, and confirmed the Y-linkage of this trait, as observed in other studies using guppies (Reynolds and Gross 1992; Reznick et al. 1997). Microsatellite analysis demonstrated that the selection regime did not affect neutral genetic diversity and to my knowledge, this work is the first study that has monitored molecular genetic variation over so few generations within selection lines. Although drift between lines was observed, no inbreeding or genetic erosion occurred over the generations bred in captivity, neither in selection lines, nor in the random breeding control (Chapter 3).

In contrast to the observation at neutral loci, we detected strong signatures of selection at individual candidate loci for male standard length, including three loci known to be in close linkage with a major QTL for male body length on the Y-chromosome (two of which have been assigned a putative function), the prolactin gene and an unmapped microsatellite marker. Both divergence in allele frequencies between treatments, as well as associations of standard length with body size were observed at these loci.

In addition, an association between standard length and two individual SNPs (M30\_Dreyer, M1046\_Dreyer) corroborated the results found elsewhere (Tripathi et al. 2009a). Tripathi et al. (2009a) used crosses between a high predation population from the Cumaná River in Venezuela and a low predation population from the Quare River from the Oropuche Drainage on Trinidad. Both these populations are from phylogenetic clades that are distinct to those used here (Willing et al. 2010). Hence, observing an association between common SNPs and standard length yields particularly strong support for their role in male standard length.

We present the first study demonstrating that fisheries-induced selection can be detected at individual loci over few generations of size-selective harvesting, revealing a quantifiable genetic basis to shifts in standard length (Fig. 4.7; 4.8). Other studies have identified QTLs for body size (e.g. Tripathi et al. 2009a), but have not analysed the adaptive evolutionary potential of this genetic variation in artificial selection experiments. Furthermore, due to the simplified nature of our experiment, other forms of selection affecting body size could be ruled out to induce patterns between the differentially harvested lines, (i.e. density-dependence, temperature). Thus, we have shown that rapid fisheries-induced evolution (FIE) can occur and that genetic change is detectable under experimental conditions.

The magnitude of the phenotypic response observed is comparable to the response observed in a similar experiment (Conover & Munch 2002) and less than five times as high as those estimated for human-harvested systems estimated by Darimont et al. (2009). What can these results say about the likelihood of genetic change by fisheries on exploited populations?

## **5.2 BROADER RELEVANCE OF THE OBTAINED RESULTS**

### **5.2.1 Life history evolution of Trinidadian guppies**

Here we have for the first time mimicked selection on body size (known to be a major driver of differentiation of wild guppy populations) under controlled conditions. In doing so, we elucidated the genetic basis of this important trait in guppy biology, and contribute to the understanding of genetic structuring among wild guppy populations exposed to contrasting predation regimes.

Assuming that the lines are representative of the wild Tacarigua population, these results suggest that both additive and non-additive genetic variation is present for standard length in an environment where this trait is under different selection pressures by predators and female mate choice. The data indicate that the genetic basis to male standard length consists prominently of Y-linked variation and furthermore that large size is associated with dominance variation and heterozygosity at underlying loci (Chapter 4).

It would be desirable and highly relevant to investigate the markers for which an association was found here in wild populations, to see whether our findings correspond with genetic variation between guppies in high- and low predation habitats. This could contribute significantly to our understanding of the role of the observed associations in determining male size in guppies, as well as of different genetic backgrounds, environmental variation and selection pressures on genetic variation in the wild at candidate genes for body size (Korsten et al. 2010). For example, no previous studies using microsatellite locus Pr39 in population genetic analyses, which exhibited strong signatures of selection in our study, have reported either significant deviations from Hardy-Weinberg equilibrium at this locus or an association between one specific allele and predation regime (e.g. van Oosterhout et al. 2006a; Suk & Neff 2009). Although these studies did not examine a relationship between Pr39 and standard length directly, a lack of evidence for selection at Pr39 in the wild may suggest that the association observed in the present study is not indicative of very close linkage of Pr39 to a major gene underlying male standard

length. Alternatively, differences in the demographic history of wild populations, together with potential gene flow (Crispo et al. 2006; Barson et al. 2009) and the polygenic basis of quantitative traits may obscure signals of divergence at single genes in the wild. Nevertheless, it might prove informative to explore the dynamics and patterns of divergence at Pr39, and closely linked loci, to investigate its relationship to variance in male standard length further.

### **5.2.2 Fisheries-induced evolution**

Despite demonstrating a direct link between harvesting selection of known intensity and measurable responses at phenotypic and genetic levels, extension of such patterns to the wild requires careful consideration. What remains to be solved, is whether current harvesting selection on commercial species is of such magnitude that it likely provokes a genetic response associated with the observed phenotypic shifts. Such a genetic response needs to be quantified and better understood before any claims on its importance for management of harvested fish species can be made (Kuparinen and Merilä 2007; Law 2007; Morita et al. 2009).

Brown et al. (2008) used data from Conover and Munch (2002) to convert the intense selection over few generations there executed to more realistic harvesting pressures for natural populations. They inferred that, for realistic commercial fishing selection pressures, a similar magnitude of evolutionary change as observed in the selection experiments by Conover and Munch (2002) could be expected over 30 generations. Such a simulation study has not been performed on our own data but nevertheless allows for coarse extrapolation of our data to more realistic scenarios. Considering that the rates of evolutionary change observed by Conover and Munch (2002) were 2-3 times our own (95-156 compared to  $50-55 \times 10^3$  d), and their selection intensity was 2.5 times as high ( $i = 0.7$  compared to  $i = 1.8$ , Falconer and Mackay 1996), a direct comparison between these two experiments is possible. Taking the extrapolation further, based on the findings by Brown et al. (2008), the magnitude of our results could be realistic for commercially exploited species over twelve generations. However, similar to the experimental approach here and in Conover and Munch (2002), Brown et al. (2008) modelled selection on discrete

generations, which is not realistic for most wild fishery scenarios. Therefore, though any response critically depends on the rate and magnitude of selection (Law 2000), a response in the wild is likely to be considerably slower than observed here (Andersen and Brander 2009). Additionally, three separate arguments can be proposed to counter such a direct comparison between our experiments and other work.

Firstly, we may have observed a greater response to selection as a result of different heritabilities. We observed  $h^2$  estimates ranging from 0.20-0.27 for selected lines and 0.52 for the control line, assuming full Y-linkage of the traits (thus a minimum  $h^2$  estimate). Full Y-linkage of standard length may be unlikely but considering the lack of a response observed in females (Chapter 2.4.3) and the  $>1$  estimate in the control line if no Y-linkage is taken into account (Chapter 3.3.5), the true  $h^2$  for our selection lines is most likely closer to  $h^2 = 0.20-0.27$  than to  $h^2 = 0.40-0.54$  (which assumes autosomal inheritance). Estimates observed here are thus comparable to those observed for wild fish species ( $h^2 = 0.2-0.5$ , and depending on whether standard length is considered as a morphological or life history trait, Weigensberg and Roff 1996; Law 2000), or those obtained by Conover and Munch ( $h^2 = 0.2$ ).

A second argument which may confound the observed strong response to selection may lie in the considerable differences in population size used here, by Conover and Munch (2002), and those realistic for wild populations. Firstly, greater population size will yield greater power for statistical tests to detect signals of selection and population differentiation (Ryman et al. 2006). Furthermore, larger populations, with larger effective population sizes and greater genetic variance, may be expected to have higher additive genetic variance at quantitative traits for selection to act upon (Hanrahan et al. 1973). The reduction in Allelic Richness but not in heterozygosity observed in our lines when comparing the F0 and F1 to the lab-born generations (see 3.3.3) indicates that such a reduction in genetic variation compared to the wild may indeed be present in our selection lines.

Of greater concern for the generality of the present study is the apparent high proportion of Y-linked genetic variation observed for male body size in guppies, and the strong influence of drift that has affected genetic differentiation between lines. Whilst the phenotypic change observed here seems comparable to other scenarios, I question whether the genetic change observed is equally analogous to the genetic basis of FIE in the wild. Both Y-linkage (no response to selection was observed in females) and drift (AMOVA  $F_{SC}$  0.033-0.034 in the F6 at microsatellites and candidate genes respectively) will have contributed to the rate of change observed in our genetic data and are unlikely to be of similar importance in larger populations, or in other species. In guppies, sexual dimorphism in morphology is particularly pronounced and male fish virtually stop growth after maturation. Such differences are less prominent or not present for Atlantic silverside (*Menidia menidia*) and many exploited species, where indeterminate growth is more common, although females often represent the larger sex as well. Consequentially, the genetic basis of body size could differ considerably across species and in other taxa, physical linkage between (sexually dimorphic) genes may be less pronounced than in guppies. In addition, other selection pressures and environmental factors in the wild may prevent or obscure efficient selection by fisheries, as has been elaborately discussed in the introduction (1.2.4 – 1.2.6).

However, the prolonged and unusually high fishing mortalities observed for commercially exploited species are unrivalled by natural selection events on adult fish (Jackson 2001; Darimont et al. 2009; Stenseth and Dunlop 2009). Large effective population sizes, which minimise effects of drift, in combination with microgeographic differentiation in some marine fish species (Conover et al. 2006; Hauser and Carvalho 2008) promote local adaptation to harvesting patterns and significant evolutionary rates by fishing. For example, Jakobsdóttir et al. (2011) observed an 80% decline in Pan I<sup>B</sup> frequency in Icelandic cod, associated with late maturing fish, over a period of 60 years. Assuming a generation time for cod of 4.8 years (Árnason et al. 2009), this corresponds to roughly 12.5 generations. Thus, genetic change in the wild under realistic harvesting pressures is possible over



similar timescales as found for the present study, strengthening confidence in the wider significance of our data.

### **5.3 FUTURE DIRECTIONS**

The findings from this study are, besides being of intrinsic value for evolutionary biology and understanding genetic dynamics on brief timescales, relevant for guppy biology, FIE and the aquaculture industry. Results on the magnitude of a selective response on both a phenotypic and a genetic level are valuable for understanding life history evolution of guppies in the wild and more broadly for evolutionary effects of fishing. Despite the simplicity of the design, the principles shown here hold in light of FIE and prove detectable genetic change can be induced at specific loci over few generations under intense directional size-specific selection. The results here obtained may be taken forward in several directions to obtain more valuable information.

#### **5.3.1 Characterisation of genetic differences between selection lines**

The complementary analyses performed in the present study have revealed genetic differentiation at candidate loci between selection lines. However, in order to fully implement the hierarchical structure of our data of replicate lines within treatments over multiple generations, more complex modelling of the data would be needed. Mixed modelling of standard length as a function of line nested within treatment and genotype as a mediator, furthermore implementing interaction effects between loci and the non-independence of data from different generations can be expected to increase the resolution of the data and provide a sounder statistical framework to distinguish between loci under selection and drift. Secondly, it is important to determine the open reading frames (ORFs) of the genes at which selection was observed. ORFs were not published with the sequences of most genes studied here. Moreover, ORFs could not be determined due to substantial unreadable portions within our own sequences, resulting from the direct sequencing of genomic fragments rather than cloned amplicons. Therefore, we have been unable to establish whether mutations observed were synonymous or non-synonymous and could not determine their function within the gene. Cloning of the desired gene

fragments to obtain whole-length readable sequences would allow for determining the most likely ORFs in our sequences.

Secondly, the number of candidate genes explored here is limited, especially considering the likely broad genetic basis of the quantitative traits examined (body size, size and age at maturation). It will therefore be extremely valuable to increase the number of markers compared between selection treatments to allow for association mapping of genes responding to the selection treatments more precisely on the guppy genome. Already, effort is being put in a follow-up project to this work, where crosses between differently selected lines will be made and segregating markers in the F2 identified using a combination of QTL linkage mapping and RAD tagging (Baird et al. 2008). Such efforts will explore more directly the role in size determination of the polymorphisms identified in the present work and will complement our findings with further candidate markers.

In addition, as already mentioned briefly, it would be very valuable to test whether signatures of selection observed in the selection lines used, are present at the same loci in the wild. Both confirming the role of these loci and the absence of an association in the wild source population or other populations, would provide valuable information on the genetic basis of standard length and generality of evolutionary pathways selecting on this important trait (Korsten et al. 2010; Rand et al. 2010).

### **5.3.2 Other putative traits under selection**

Because only adult size was measured in the present study and, with exception of generation F2, not growth rate, it was not possible to infer whether the observed differences between selection lines were the result of differences in growth rate besides absolute growth. Obtaining a quantification of growth rate and differences herein between lines would be highly relevant, as a substantial part of the debate on FIE is centred around its effects on growth (e.g. Hilborn and Minte-Vera 2008; Kuparinen et al. 2008; Nusslé et al. 2009).

Secondly, there is considerable evidence suggesting correlations between life history traits and other traits, such as aspects of morphology and behaviour (Ruzzante and Doyle 1991; Walsh et al. 2006; Chiba et al. 2007; Biro and Stamps 2008; Uusi-Heikkilä et al. 2008). Quantifying a correlated response to selection on body size of life history traits (maturation, reproductive allotment, reproductive output), behaviour (mating preference, schooling, personality traits) and morphology (colour patterns, morphometrics) would be extremely interesting. Both the field of guppy biology and fisheries science would benefit from an improved understanding of relationships between such traits.

Finally, it would be interesting to investigate the response to relaxing and/ or reversing the selection intensity. Measuring the response to a change in the selection regime could yield insights into the different pathways the response to selection can take: Does sufficient additive genetic variance remain for selection in the opposite direction? Are rates of evolutionary change comparable between the initial response and the response to reverse selection? Does selection act upon the same genes when its direction is reversed or is previously untouched variation utilised? To be able to answer such questions would be highly relevant for understanding effects of FIE and how these are best incorporated into management strategies (Conover et al. 2009).

### **5.3.3 The significance of evolutionary principles for fisheries management**

We confirm that genetic change by size selective harvesting can occur on contemporary time scales and these experimental results are indirectly comparable to the effects of commercial fishing in the wild. Thereby, this study provides support for the need to incorporate evolutionary principles into fisheries management (Stokes and Law 2000; Jorgensen et al. 2007; Dunlop et al. 2009). Although our results do not provide insights into the likely reversibility of a genetic response to size selective harvesting, it has been shown experimentally that selective harvesting over few generations can be reversed (Conover et al. 2009). However, modelling studies suggest this can be expected to be a slow process in the wild (Enberg et al. 2009) and a precautionary principle towards fisheries-induced genetic change seems

justified (Law 2007; Dunlop et al. 2009). “Fish less and fish more carefully, protecting the larger individuals of the stocks”, is a message heard increasingly (Stokes and Law 2000; Law 2007; Hilborn and Minto-Vera 2008; Dunlop et al. 2009). However, considering the first advice of this kind was proposed at the beginning of the last century (Rutter 1902) current managers seem to lack either the insight or the power to enforce such changes in management strategies. While debate on the importance of FIE for sustaining fisheries yield and high quality stocks has placed the topic higher on the political and commercial agenda (Browman et al. 2000; Jorgensen et al. 2007; Dunlop et al. 2009), ambiguity within the scientific community has not facilitated change in management strategies. I hope the present study will contribute to a better understanding of the genetic effects of selective fishing and provides a constructive element for the development of sustainable fisheries.

## **REFERENCES**

- Aitken, N., Smith, S., Schwarz, C., Morin, P.A., 2004. Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. *Molecular Ecology* 13: 1423-1431.
- Alexander, H.J., Taylor, J.S., Wu, S.S., Breden, F., 2006. Parallel evolution and vicariance in the guppy (*Poecilia reticulata*) over multiple spatial and temporal scales. *Evolution* 60: 2352–2369.
- Allendorf, F.W., England, P., Luikart, G., Ritchie, P., Ryman, N., 2008. Genetic effects of harvest on wild animal populations. *Trends in Ecology & Evolution* 23: 327-337.
- Almuly, R., Poleg-Danin, Y., Gorshkov, S., Gorshikova, G., Rapoport, B., Soller, M., Kashi, Y., Funkenstein, B., 2005. Characterization of the 5' flanking region of the growth hormone gene of the marine teleost, gilthead sea bream *Sparus aurata*: analysis of a polymorphic microsatellite in the proximal promoter. *Fisheries Science* 71: 479-490.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* 25: 3389.
- Andersen, K.H., Brander, K., 2009. Expected rate of fisheries-induced evolution is slow. *Proceedings of the National Academy of Sciences* 106: 11657-11660.
- Antao, T., Lopes, A., Lopes, R., Beja-Pereira, A., Luikart, G., 2008. LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC bioinformatics* 9: 323.
- Arendt, J.D., Reznick, D.N., 2005. Evolution of juvenile growth rates in female guppies (*Poecilia reticulata*): predator regime or resource level? *Proceedings of the Royal Society B: Biological Sciences* 272: 333-337.
- Árnason, E., Hernandez, U.B., Kristinsson, K., 2009. Intense habitat-specific fisheries-induced selection at the molecular Pan I locus predicts imminent collapse of a major cod fishery. *PLoS ONE* 4: e5529.
- Arreguín-Sánchez, F., 1996. Catchability: a key parameter for fish stock assessment. *Reviews in Fish Biology and Fisheries* 6: 221–242.
- Atchley, W.R., Fitch, W.M., Bronner-Fraser, M., 1994. Molecular evolution of the MyoD family of transcription factors. *Proceedings of the National Academy of Sciences* 91: 11522-11526.
- Avilés, J.M., Stokke, B.G., Moksnes, A., Roskaft, E., Asmul, M., Moller, A.P., 2006. Rapid increase in cuckoo egg matching in a recently parasitized reed warbler population. *Journal of Evolutionary Biology* 19: 1901–1910.
- Avise, J.C., 2004. *Molecular Markers, Natural History, and Evolution*, 2nd ed. Sinauer Associates, Sunderland, MA.
- Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3: e3376.
- Barbosa, M., Magurran, A.E., 2006. Female mating decisions: maximizing fitness? *Journal of Fish Biology* 68: 1636-1661.
- Barot, S., Heino, M., Morgan, M.J., Dieckmann, U., 2005. Maturation of Newfoundland American plaice (*Hippoglossoides platessoides*): long-term Trends in maturation reaction norms despite low fishing mortality? *ICES Journal of Marine Science* 62: 56-64.

- Barrett, R.D.H., Paccard, A., Healy, T.M., Bergek, S., Schulte, P.M., Schluter, D., Rogers, S.M., 2010. Rapid evolution of cold tolerance in stickleback. *Proceedings of the Royal Society B: Biological Sciences* 278: 233-238.
- Barson, N.J., Cable, J., van Oosterhout, C., 2009. Population genetic analysis of microsatellite variation of guppies (*Poecilia reticulata*) in Trinidad and Tobago: evidence for a dynamic source–sink metapopulation structure, founder events and population bottlenecks. *Journal of Evolutionary Biology* 22: 485–497.
- Baulier, L., Heino, M., Lilly, G.R., Dieckmann, U., 2006. Body condition and evolution of maturation of Atlantic cod in Newfoundland. *ICES CM* 2006/H:19.
- Beaumont, M.A., Nichols, R.A., 1996. Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 263: 1619-1626.
- Becher, S.A., Magurran, A.E., 2004. Multiple mating and reproductive skew in Trinidadian guppies. *Proceedings of the Royal Society B: Biological Sciences* 271: 1009-1014.
- Becher, S.A., Russell, S.T., Magurran, A.E., 2002. Isolation and characterization of polymorphic microsatellites in the Trinidadian guppy (*Poecilia reticulata*). *Molecular Ecology Notes* 2: 456-458.
- Bekkevold, D., Hansen, M.M., Loeschcke, V., 2002. Male reproductive competition in spawning aggregations of cod (*Gadus morhua*, L.). *Molecular Ecology* 11: 91–102.
- Beldade, P.Í., Koops, K., Brakefield, P.M., 2002. Developmental constraints versus flexibility in morphological evolution. *Nature* 416: 844–847.
- Bell, M.A., Aguirre, W.E., Buck, N.J., 2004. Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution* 58: 814–824.
- Belter, J.G., Carey, H.V., Garland, T., 2004. Effects of voluntary exercise and genetic selection for high activity levels on HSP72 expression in house mice. *Journal of Applied Physiology* 96: 1270.
- Beraldi, D., McRae, A.F., Gratten, J., Slate, J., Visscher, P.M., Pemberton, J.M., 2006. Development of a Linkage Map and mapping of phenotypic polymorphisms in a free-living population of Soay Sheep (*Ovis aries*). *Genetics* 173: 1521-1537.
- Berkeley, S.A., Hixon, M.A., Larson, R.J., Love, M.S., 2004. Fisheries sustainability via protection of age structure and spatial distribution of fish populations. *Fisheries* 29: 23–32.
- Bernal-Ramírez, J.H., Adcock, G.J., Hauser, L., Carvalho, G.R., Smith, P.J., 2003. Temporal stability of genetic population structure in the New Zealand snapper, *Pagrus auratus*, and relationship to coastal currents. *Marine Biology* 142: 567-574.
- Bernatchez, L., Landry, C., 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology* 16: 363–377.
- Beverton, R.J.H., 1990. Small marine pelagic fish and the threat of fishing; are they endangered? *Journal of Fish Biology* 37: 5–16.
- Beverton, R.J.H., 1998. Fish, fact and fantasy: a long view. *Reviews in Fish Biology and Fisheries* 8: 229–249.
- Bigler, B.S., Welch, D.W., Helle, J.H., 1996. A review of size Trends among North Pacific salmon (*Oncorhynchus spp.*). *Canadian Journal of Fisheries and Aquatic Sciences* 53: 455–465.

- Biro, P.A., Post, J.R., 2008. Rapid depletion of genotypes with fast growth and bold personality traits from harvested fish populations. *Proceedings of the National Academy of Sciences* 105: 2919.
- Biro, P.A., Stamps, J.A., 2008. Are animal personality traits linked to life-history productivity? *Trends in Ecology & Evolution* 23: 361-368.
- Blel, H., Panfili, J., Guinand, B., Berrebi, P., Said, K., Durand, J.-D., 2010. Selection footprint at the first intron of the Prl gene in natural populations of the flathead mullet (*Mugil cephalus*, L. 1758). *Journal of Experimental Marine Biology and Ecology* 387: 60-67.
- Bole-Feyssot, C., Goffin, V., Edery, M., Binart, N., Kelly, P.A., 1998. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. *Endocrine Reviews* 19: 225-268.
- Bonin, A., Taberlet, P., Miaud, C., Pompanon, F., 2006. Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution* 23: 773-783.
- Botsford, L.W., 1997. The Management of Fisheries and Marine Ecosystems. *Science* 277: 509-515.
- Bradley, R.D., Hillis, D.M., 1997. Recombinant DNA sequences generated by PCR amplification. *Molecular Biology and Evolution* 14: 592.
- Brakefield, P.M., 2003. Artificial selection and the development of ecologically relevant phenotypes. *Ecology* 84: 1661-1671.
- Brander, K.M., 2007. The role of growth changes in the decline and recovery of North Atlantic cod stocks since 1970. *ICES Journal of Marine Science* 64: 211-217.
- Breden, F., Ptacek, M.B., Rashed, M., Taphorn, D., Figueiredo, C.A., 1999. Molecular phylogeny of the live-bearing fish genus *Poecilia* (Cyprinodontiformes: Poeciliidae). *Molecular Phylogenetics and Evolution* 12: 95-104.
- Brooks, R., 2000. Negative genetic correlation between male sexual attractiveness and survival. *Nature* 406: 67-70.
- Browman, H.I., Browman, H.I., Hutchings, J.A., Conover, D.O., Stokes, K., Law, R., Walters, C., 2000. Evolution' of fisheries science. *Marine Ecology Progress Series* 208: 299-313.
- Brown, J.H., Marquet, P.A., Taper, M.L., 1993. Evolution of body size: consequences of an energetic definition of fitness. *The American Naturalist* 142: 573-584.
- Brown, C.J., Hobday, A.J., Ziegler, P.E., Welsford, D.C., 2008. Darwinian fisheries science needs to consider realistic fishing pressures over evolutionary time scales. *Marine Ecology Progress Series* 369: 257-266.
- Bulmer, M.G., 1971. The effect of selection on genetic variability. *The American Naturalist* 105: 201-211.
- Burke, M.K., Dunham, J.P., Shahrestani, P., Thornton, K.R., Rose, M.R., Long, A.D., 2010. Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* 467: 587-590.
- Butlin, R.K., 2008. Population genomics and speciation. *Genetica* 138: 409-418.
- Cable, J., van Oosterhout, C., 2007. The role of innate and acquired resistance in two natural populations of guppies (*Poecilia reticulata*) infected with the ectoparasite *Gyrodactylus turnbulli*. *Biological Journal of the Linnean Society* 90: 647-655.
- Carroll, S.P., Hendry, A.P., Reznick, D.N., Fox, C.W., 2007. Evolution on ecological time-scales. *Functional Ecology* 21: 387-393.



- Carvajal-Rodríguez, A., de Uña-Alvarez, J., Rolán-Alvarez, E., 2009. A new multitest correction (SGoF) that increases its statistical power when increasing the number of tests. *BMC bioinformatics* 10: 209.
- Carvalho, G.R., van Oosterhout, C., Hauser, L., Magurran, A.E., 2003. Measuring genetic variation in wild populations: from molecular markers to adaptive traits. In: Hails, R.S., Beringer, J.E., Godfray, H.J.C. (Eds.), *Genes in the Environment*. Blackwell Publishing, Oxford, pp. 91-111.
- Carvalho, G.R., Pitcher, T.J., 1994. *Molecular genetics in fisheries*. Chapman and Hall, London.
- Carvalho, G.R., Shaw, P.W., Hauser, L., Seghers, B.H., Magurran, A.E., 1996. Artificial introductions, evolutionary change and population differentiation in Trinidadian guppies (*Poecilia reticulata*; Poeciliidae). *Biological Journal of the Linnean Society* 57: 219–234.
- Carvalho, G.R., Shaw, P.W., Magurran, A.E., Seghers, B.H., 1991. Marked genetic divergence revealed by allozymes among populations of the guppy *Poecilia reticulata* (Poeciliidae), in Trinidad. *Biological Journal of the Linnean Society* 42: 389–405.
- Charlesworth, B., Charlesworth, D., 1999. The genetic basis of inbreeding depression. *Genetics Research Cambridge* 74: 329–340.
- Chiba, S., Arnott, S.A., Conover, D.O., 2007. Coevolution of foraging behavior with intrinsic growth rate: risk-taking in naturally and artificially selected growth genotypes of *Menidia menidia*. *Oecologia* 154: 237-246.
- Chippindale, A.K., 2006. *Experimental evolution*. In: *Evolutionary Genetics: Concepts and Case Studies*. Oxford University Press, Oxford.
- Chippindale, A.K., Alipaz, J.A., Chen, H.W., Rose, M.R., 1997. Experimental evolution of accelerated development in *Drosophila*. 1. Developmental speed and larval survival. *Evolution* 51: 1536–1551.
- Coltman, D.W., 2008. Evolutionary rebound from selective harvesting. *Trends in Ecology & Evolution* 23: 117–118.
- Coltman, D.W., O’Donoghue, P., Jorgenson, J.T., Hogg, J.T., Strobeck, C., Festa-Bianchet, M., 2003. Undesirable evolutionary consequences of trophy hunting. *Nature* 426: 655–658.
- Conover, D.O., 2000. Darwinian fishery science. *Marine Ecology Progress Series* 208: 303-307.
- Conover, D.O., 2007. Nets versus nature. *Nature* 450: 179–180.
- Conover, D.O., Baumann, H., 2009. The role of experiments in understanding fishery-induced evolution. *Evolutionary Applications* 2: 276–290.
- Conover, D.O., Clarke, L.M., Munch, S.B., Wagner, G.N., 2006. Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *Journal of Fish Biology* 69 (Suppl. C): 21-47.
- Conover, D.O., Munch, S.B., 2002. Sustaining fisheries yields over evolutionary time scales. *Science* 297: 94.
- Conover, D.O., Munch, S.B., Arnott, S.A., 2009. Reversal of evolutionary downsizing caused by selective harvest of large fish. *Proceedings of the Royal Society B: Biological Sciences* 276: 2015-2020.

- Conover, D.O., Present, T.M., 1990. Countergradient variation in growth rate: compensation for length of the growing season among Atlantic silversides from different latitudes. *Oecologia* 83: 316–324.
- Conover, D.O., Schultz, E.T., 1995. Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution* 10: 248-252.
- Courtenay, W.R., Meffe, G.K., 1989. Small fish in strange places: a review of introduced poeciliids. In: *Ecology and Evolution of Livebearing Fishes*. Prentice Hall, Englewood Cliffs, NJ, pp. 319-331.
- Crispo, E., Bentzen, P., Reznick, D.N., Kinnison, M.T., Hendry, A.P., 2006. The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology* 15: 49–62.
- Croft, D.P., Albanese, B., Arrowsmith, B.J., Botham, M., Webster, M., Krause, J., 2003a. Sex-biased movement in the guppy (*Poecilia reticulata*). *Oecologia* 137: 62–68.
- Croft, D.P., Arrowsmith, B.J., Bielby, J., Skinner, K., White, E., Couzin, I.D., Magurran, A.E., Ramnarine, I., Krause, J., 2003b. Mechanisms underlying shoal composition in the Trinidadian guppy, *Poecilia reticulata*. *Oikos* 100: 429–438.
- Croft, D.P., Darden, S.K., Ruxton, G.D., 2009. Predation risk as a driving force for phenotypic assortment: a cross-population comparison. *Proceedings of the Royal Society B: Biological Sciences* 276: 1899-1904.
- Cuveliers, E.L., Volckaert, F.A.M., Rijnsdorp, A.D., Larmuseau, M.H.D., Maes, G.E., 2011. Temporal genetic stability and high effective population size despite fisheries-induced life-history trait evolution in the North Sea sole. *Molecular Ecology* 20: 3555-3568.
- Darimont, C.T., Carlson, S.M., Kinnison, M.T., Paquet, P.C., Reimchen, T.E., Wilmers, C.C., 2009. Human predators outpace other agents of trait change in the wild. *Proceedings of the National Academy of Sciences* 106: 952-954.
- Darwin, C., 1859. *On the Origin of Species*. John Murray, London, UK.
- Daufresne, M., Lengfellner, K., Sommer, U., 2009. Global warming benefits the small in aquatic ecosystems. *Proceedings of the National Academy of Sciences* 106: 12788-12793.
- De-Santis, C., Jerry, D.R., 2007. Candidate growth genes in finfish - Where should we be looking? *Aquaculture* 272: 22-38.
- DeWoody, J.A., Avise, J.C., 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology* 56: 461-473.
- DiBattista, J.D., Feldheim, K.A., Garant, D., Gruber, S.H., Hendry, A.P., 2009. Evolutionary potential of a large marine vertebrate: quantitative genetic parameters in a wild population. *Evolution* 63-4: 1051-1067.
- Dieckmann, U., Heino, M., 2007. Probabilistic maturation reaction norms: their history, strengths, and limitations. *Marine Ecology Progress Series* 335: 253-269.
- Douglas, S., Knickle, L., Williams, J., Flight, R., Reith, M., 2008. A first generation Atlantic halibut *Hippoglossus hippoglossus* (L.) microarray: application to developmental studies. *Journal of Fish Biology* 72: 2391–2406.
- Dreyer, C., Hoffmann, M., Lanz, C., Willing, E.M., Riester, M., Warthmann, N., Sprecher, A., Tripathi, N., Henz, S.R., Weigel, D., 2007. ESTs and EST-linked polymorphisms for genetic mapping and phylogenetic reconstruction in the guppy, *Poecilia reticulata*. *BMC genomics* 8: 269.

- Dulvy, N.K., Sadovy, Y., Reynolds, J.D., 2003. Extinction vulnerability in marine populations. *Fish and Fisheries* 4: 25–64.
- Dunlop, E.S., Enberg, K., Jørgensen, C., Heino, M., 2009. Toward Darwinian fisheries management. *Evolutionary Applications* 2: 245–259.
- Edeline, E., Carlson, S.M., Stige, L.C., Winfield, I.J., Fletcher, J.M., James, J.B., Haugen, T.O., Völlestad, L.A., Stenseth, N.C., 2007. Trait changes in a harvested population are driven by a dynamic tug-of-war between natural and harvest selection. *Proceedings of the National Academy of Sciences* 104: 15799.
- Edley, M.T., Law, R., 1988. Evolution of life histories and yields in experimental populations of *Daphnia magna*. *Biological Journal of the Linnean Society* 34: 309–326.
- El Mousadik, A., Petit, R.J., 1996. High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. *Theoretical and Applied Genetics* 92: 832–839.
- Ellegren, H., Sheldon, B.C., 2008. Genetic basis of fitness differences in natural populations. *Nature* 452: 169–175.
- Enberg, K., Jørgensen, C., Dunlop, E.S., Heino, M., Dieckmann, U., 2009. Implications of fisheries-induced evolution for stock rebuilding and recovery. *Evolutionary Applications* 2: 394–414.
- Endler, J.A., 1978. A predator's view of animal color patterns. *Evolutionary biology* 11: 319–364.
- Endler, J.A., 1980. Natural selection on color patterns in *Poecilia reticulata*. *Evolution* 34: 76–91.
- Endler, J.A., 1983. Natural and sexual selection on color patterns in poeciliid fishes. *Environmental Biology of Fishes* 9: 173–190.
- Endler, J.A., 1986. *Natural Selection in the Wild*, Monographs in Population Biology. Princeton University Press, Princeton, NJ.
- Endler, J.A., 1987. Predation, light intensity and courtship behaviour in *Poecilia reticulata* (Pisces: Poeciliidae). *Animal Behaviour* 35: 1376–1385.
- Endler, J.A., 1988. Sexual selection and predation risk in guppies. *Nature* 332: 593–594.
- Endler, J.A., 1995. Multiple-trait coevolution and environmental gradients in guppies. *Trends in Ecology & Evolution* 10: 22–29.
- Endler, J.A., Houde, A.E., 1995. Geographic variation in female preferences for male traits in *Poecilia reticulata*. *Evolution* 49: 456–468.
- Engelhard, G.H., Heino, M., 2004a. Maturity changes in Norwegian spring-spawning herring *Clupea harengus*: compensatory or evolutionary responses? *Marine ecology progress series* 272: 245–256.
- Engelhard, G.H., Heino, M., 2004b. Maturity changes in Norwegian spring-spawning herring before, during, and after a major population collapse. *ICES Journal of Marine Science* 60: 304–313.
- England, P.R., Osler, G.H., Woodworth, L.M., Montgomery, M.E., Briscoe, D.A., Frankham, R., 2003. Effects of intense versus diffuse population bottlenecks on microsatellite genetic diversity and evolutionary potential. *Conservation Genetics* 4: 595–604.
- Ernande, B., Dieckmann, U., Heino, M., 2004. Adaptive changes in harvested populations: plasticity and evolution of age and size at maturation. *Proceedings of the Royal Society B: Biological Sciences* 271: 415–423.

- Estes, J.A., Tinker, M.T., Williams, T.M., Doak, D.F., 1998. Killer whale predation on sea otters linking oceanic and nearshore ecosystems. *Science* 282: 473.
- Evans, J.P., Magurran, A.E., 2000. Multiple benefits of multiple mating in guppies. *Proceedings of the National Academy of Sciences of the United States of America* 97: 10074.
- Evans, J.P., Magurran, A.E., 2001. Patterns of sperm precedence and predictors of paternity in the Trinidadian guppy. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268: 719-724.
- Evans, J.P., Zane, L., Francescato, S., Pilastro, A., 2003. Directional postcopulatory sexual selection revealed by artificial insemination. *Nature* 421: 360-363.
- Excoffier, L., Hofer, T., Foll, M., 2009. Detecting loci under selection in a hierarchically structured population. *Heredity* 103: 285-298.
- Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47-50.
- FAO, 2010. *The State of World Fisheries and Aquaculture*. Food and Agriculture Organization of the United Nations, Rome.
- Fajen, A., Breden, F., 1992. Mitochondrial DNA Sequence Variation among Natural Populations of the Trinidadian Guppy, *Poecilia reticulata*. *Evolution* 46: 1457-1465.
- Falconer, D.S., 1953. Selection for large and small size in mice. *Journal of Genetics* 51: 470-501.
- Falconer, D.S., Mackay, T.F.C., 1996. *Introduction to Quantitative Genetics*, 4th ed. Longman, Harlow.
- Fenberg, P.B., Roy, K., 2008. Ecological and evolutionary consequences of size-selective harvesting: how much do we know? *Molecular Ecology* 17: 209-220.
- Fernö, A., Olsen, S., 1994. *Marine fish behaviour in capture and abundance estimation*. Fishing News Books, Oxford.
- Fidler, A.E., van Oers, K., Drent, P.J., Kuhn, S., Mueller, J.C., Kempenaers, B., 2007. Drd4 gene polymorphisms are associated with personality variation in a passerine bird. *Proceedings of the Royal Society B: Biological Sciences* 274: 1685.
- Fisher, R.A., 1930. *The genetical theory of natural selection*. Oxford University Press, Oxford.
- Fitzpatrick, M., Bensch, Y., Smid, H., Vet, L., Robinson, G., Sokolowski, M., 2005. Candidate genes for behavioural ecology. *Trends in Ecology & Evolution* 20: 96-104.
- Florini, J.R., Ewton, D.Z., Coolican, S.A., 1996. Growth hormone and the Insulin-like Growth Factor system in myogenesis. *Endocrine Reviews* 17: 481-517.
- Fogarty, M.J., Sissenwine, M.P., Cohen, E.B., 1991. Recruitment Variability and the Dynamics of Exploited Populations. *Trends in Ecology & Evolution* 6: 241-246.
- Fraser, B.A., Ramnarine, I.W., Neff, B.D., 2009. Selection at the MHC class IIB locus across guppy (*Poecilia reticulata*) populations. *Heredity* 104: 155-167.
- Fraser, B.A., Ramnarine, I.W., Neff, B.D., 2010. Temporal variation at the MHC class IIB in wild populations of the guppy (*Poecilia reticulata*). *Evolution* 64: 2086-2096.
- Fraser, B.A., Weadick, C.J., Janowitz, I., Rodd, F.H., Hughes, K.A., 2011. Sequencing and characterization of the guppy (*Poecilia reticulata*) transcriptome. *BMC genomics* 12.
- Fraser, D.F., Gilliam, J.F., Yip-Hoi, T., 1995. Predation as an agent of population fragmentation in a tropical watershed. *Ecology* 76: 1461-1472.

- Fuller, R.C., Baer, C.F., Travis, J., 2005. How and when selection experiments might actually be useful. *Integrative and Comparative Biology* 45: 391.
- Funk, W.C., Tyburczy, J.A., Knudsen, K.L., Lindner, K.R., Allendorf, F.W., 2005. Genetic basis of variation in morphological and life-history traits of a wild population of pink salmon. *Journal of Heredity* 96: 24.
- Garant, D., Kruuk, L.E., Wilkin, T.A., McCleery, R.H., Sheldon, B.C., 2005. Evolution driven by differential dispersal within a wild bird population. *Nature* 433: 60–65.
- Garant, D., Kruuk, L.E.B., 2005. How to use molecular marker data to measure evolutionary parameters in wild populations. *Molecular Ecology* 14: 1843–1859.
- Garland Jr., T., Morgan, M.T., Swallow, J.G., Rhodes, J.S., Girard, I., Belter, J.G., Carter, P.A., 2002. Evolution of a small-muscle polymorphism in lines of house mice selected for high activity levels. *Evolution* 56: 1267.
- Ghalambor, C.K., McKay, J.K., Carroll, S.P., Reznick, D.N., 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* 21: 394-407.
- Gienapp, P., Teplitsky, C., Alho, J.S., Mills, J.A., Merila, J., 2008. Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology* 17: 167–178.
- Gingerich, P.D., 1983. Rates of evolution: effects of time and temporal scaling. *Science* 222: 159-161.
- Godin, J.-G.J., Briggs, S.E., 1996. Female mate choice under predation risk in the guppy. *Animal Behaviour* 51: 117-130.
- Gordon, S.P., Reznick, D.N., Kinnison, M.T., Bryant, M. J., Weese, D.J., Räsänen, K., Millar, N.P., Hendry, A.P., 2009. Adaptive changes in life history and survival following a new guppy introduction. *The American Naturalist* 174: 34-45.
- Gorman, K.F., Christians, J.K., Parent, J., Ahmadi, R., Weigel, D., Dreyer, C., Breden, F., 2011. A major QTL controls susceptibility to spinal curvature in the curveback guppy. *BMC genetics* 12: 16.
- Goudet, J., 1995. FSTAT (vers. 1.2): a computer program to calculate F-statistics. *Journal of Heredity* 86: 486-486.
- Grant, P.R., Grant, B.R., 2002. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296: 707-711.
- Gratten, J., Beraldi, D., Lowder, B., McRae, A., Visscher, P., Pemberton, J., Slate, J., 2007. Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings of the Royal Society B: Biological Sciences* 274: 619-626.
- Gratten, J., Wilson, A.J., McRae, A.F., Beraldi, D., Visscher, P.M., Pemberton, J.M., Slate, J., 2008. A localized negative genetic constrains microevolution of coat color in wild sheep. *Science* 319: 318-320.
- Grether, G.F., Millie, D.F., Bryant, M.J., Reznick, D.N., Mayea, W., 2001. Rain forest canopy cover, resource availability, and life history evolution in guppies. *Ecology* 82: 1546–1559.
- Grift, R.E., Heino, M., Rijnsdorp, A.D., Kraak, S.B., Dieckmann, U., 2007. Three-dimensional maturation reaction norms for North Sea plaice. *Marine Ecology Progress Series* 334: 213–224.

- Grift, R.E., Rijnsdorp, A.D., Barot, S., Heino, M., Dieckmann, U., 2003. Fisheries-induced trends in reaction norms for maturation in North Sea plaice. *Marine Ecology Progress Series* 257: 247–257.
- Günther, A.C.L.G., 1866. *Catalogue of the fishes in the British Museum*. British Museum (Natural History).
- Handford, P., Bell, G., Reimchen, T., 1977. A gillnet fishery considered as an experiment in artificial selection. *Journal of the Fisheries Research Board of Canada* 34: 954-961.
- Hanrahan, J.P., Eisen, E.J., Legates, J.E., 1973. Effects of population size and selection intensity on short-term response to selection for postweaning gain in mice. *Genetics* 73: 513.
- Hansen, M.M., Kenchington, E., Nielsen, E.E., 2001. Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries* 2: 93–112.
- Hansen, M.M., Meier, K., Mensberg, K.L.D., 2010. Identifying footprints of selection in stocked brown trout populations: a spatio-temporal approach. *Molecular Ecology* 19: 1787–1800.
- Hard, J.J., Gross, M.R., Heino, M., Hilborn, R., Kope, R.G., Law, R., Reynolds, J.D., 2008. Synthesis: Evolutionary consequences of fishing and their implications for salmon. *Evolutionary Applications* 1: 388–408.
- Hardin, G., 1968. The tragedy of the commons. *Science* 162: 1243-1248.
- Hartl, D.L., Clark, A.G., 1989. *Principles of population genetics*. Sinauer Associates, Sunderland, MA.
- Haskins, C.P., Haskins, E.F., 1949. The role of sexual selection as an isolating mechanism in three species of Poeciliid fishes. *Evolution* 3: 160-169.
- Haskins, C.P., Haskins, E.F., 1951. The inheritance of certain color patterns in wild populations of *Lebistes reticulatus* in Trinidad. *Evolution* 5: 216-225.
- Haskins, C.P., Haskins, E.F., McLaughlin, J.J.A., Hewitt, J.E., 1961. Polymorphism and population structure in *Lebistes reticulatus*, an ecological study. In: *Vertebrate Speciation*. University of Texas Press, Austin, pp. 320-395.
- Haubruege, E., Petit, F., Gage, M.J.G., 2000. Reduced sperm counts in guppies (*Poecilia reticulata*) following exposure to low levels of tributyltin and bisphenol A. *Proceedings of the Royal Society B: Biological Sciences* 267: 2333-2337.
- Haugen, T.O., Völlestad, L.A., 2001. A century of life-history evolution in grayling. *Genetica* 112-113: 475-491.
- Hauser, L., Adcock, G.J., Smith, P.J., Bernal-Ramírez, J.H., Carvalho, G.R., 2002. Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences* 99: 11742.
- Hauser, L., Carvalho, G.R., 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries* 9: 333–362.
- Hauser, L., Seeb, J.E., 2008. Advances in molecular technology and their impact on fisheries genetics. *Fish and Fisheries* 9: 473-486.
- Heap, I.M., 1997. The occurrence of herbicide-resistant weeds. *Pesticide Science* 51: 235-243.
- Hedgecock, D., 1994. Temporal and spatial genetic structure of marine animal populations in the California Current. *California Cooperative Oceanic Fisheries Investigations Reports* 35: 73-81.

- Heiman, M.L., Chen, Y., Caro, J.F., 1998. Leptin participates in the regulation of glucocorticoid and growth hormone axes<sup>1</sup>. *The Journal of Nutritional Biochemistry* 9: 553–559.
- Heino, M., 1998. Management of evolving fish stocks. *Canadian Journal of Fisheries and Aquatic Sciences* 55: 1971–1982.
- Heino, M., Dieckmann, U., Godo, O.R., 2002. Estimating reaction norms for age and size at maturation with reconstructed immature size distributions: a new technique illustrated by application to Northeast Arctic cod. *ICES Journal of Marine Science: Journal du Conseil* 59: 562.
- Heino, M., Godo, O.R., 2002. Fisheries-induced selection pressures in the context of sustainable fisheries. *Bulletin of Marine Science* 70: 639–656.
- Heino, M., Kaitala, V., 1999. Evolution of resource allocation between growth and reproduction in animals with indeterminate growth. *Journal of Evolutionary Biology* 12: 423–429.
- Hemmer-Hansen, J., Nielsen, E.E., Frydenberg, J., Loeschcke, V., 2007. Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus* L.). *Heredity* 99: 592–600.
- Hemmer-Hansen, J., Nielsen, E.E., Meldrup, D., Mittelholzer, C., 2011. Identification of single nucleotide polymorphisms in candidate genes for growth and reproduction in a nonmodel organism; the Atlantic cod, *Gadus morhua*. *Molecular Ecology Resources*.
- Hendry, A.P., 2000. Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science* 290: 516–518.
- Hendry, A.P., Farrugia, T.J., Kinnison, M.T., 2008. Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology* 17: 20–29.
- Hendry, A.P., Kinnison, M.T., 1999. The pace of modern life: measuring rates of contemporary microevolution. *Evolution* 53: 1637–1653.
- Hendry, A.P., Kinnison, M.T., 2001. An introduction to microevolution: rate, pattern, process. *Genetica* 112: 1–8.
- Hendry, A.P., Kinnison, M.T., Heino, M., Day, T., Smith, T.B., Fitt, G., Bergstrom, C.T., Oakeshott, J., Jorgensen, P., Zalucki, M.P., Gilchrist, G., Southerton, S., Sih, A., Strauss, S., Denison, R.F., Carroll, S.P., 2011. Evolutionary principles and their application. *Evolutionary Applications* 4: 159–183.
- Hepp, G.R., Kenamer, R.A., Johnson, M.H., 2006. Maternal effects in Wood Ducks: incubation temperature influences incubation period and neonate phenotype. *Functional Ecology* 20: 308–314.
- Hilborn, R., 2006. Faith-based fisheries. *Fisheries* 31: 554–555.
- Hilborn, R., Minto-Vera, C.V., 2008. Fisheries-induced changes in growth rates in marine fisheries: are they significant? *Bulletin of Marine science* 83: 95–105.
- Hilborn, R., Quinn, T.P., Schindler, D.E., Rogers, D.E., 2003. Biocomplexity and fisheries sustainability. *Proceedings of the National Academy of Sciences of the United States of America* 100: 6564.
- Hill, W.G., Caballero, A., 1992. Artificial selection experiments. *Annual Review of Ecology and Systematics* 23: 287–310.
- Hoarau, G., Boon, E., Jongma, D.N., Ferber, S., Palsson, J., van der Veer, H.W., Rijnsdorp, A.D., Stam, W.T., Olsen, J.L., 2005. Low effective population size and evidence for

- inbreeding in an overexploited flatfish, plaice (*Pleuronectes platessa* L.). Proceedings of the Royal Society B: Biological Sciences 272: 497.
- Hoffmann, A.A., Merilä, J., 1999. Heritable variation and evolution under favourable and unfavourable conditions. Trends in Ecology and Evolution 14: 96–101.
- Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A., Cresko, W.A., 2010. Population genomics of parallel adaptation in Threespine Stickleback using sequenced RAD tags. PLoS Genet 6: e1000862.
- Holloway, G.J., Povey, S.R., Sibly, R.M., 1990. The effect of new environment on adapted genetic architecture. Heredity 64: 323-330.
- Hosking, L., Lumsden, S., Lewis, K., Yeo, A., McCarthy, L., Bansal, A., Riley, J., Purvis, I., Xu, C.-F., 2004. Detection of genotyping errors by Hardy–Weinberg equilibrium testing. European Journal of Human Genetics 12: 395-399.
- Houde, A.E., 1992. Sex-linked heritability of a sexually selected character in a natural populations of *Poecilia reticulata* (Pisces: Poeciliidae) (guppies). Heredity 69: 229-235.
- Houde, A.E., 1997. Sex, Color, and Mate Choice in Guppies, Monographs in Behavior and Ecology. Princeton University Press, Princeton.
- Houle, D., 1992. Comparing evolvability and variability of quantitative traits. Genetics 130: 195.
- Houle-Leroy, P., Guderley, H., Swallow, J.G., Garland Jr., T., 2003. Artificial selection for high activity favors mighty mini-muscles in house mice. American Journal of Physiology Regulatory, Integrative and Comparative Physiology 284: R433-R443.
- Hsieh, C., Reiss, C.S., Hunter, J.R., Beddington, J.R., May, R.M., Sugihara, G., 2006. Fishing elevates variability in the abundance of exploited species. Nature 443: 859-862.
- Hutchings, J.A., 2000. Collapse and recovery of marine fishes. Nature 406: 882-885.
- Hutchings, J.A., Swain, D.P., Rowe, S., Eddington, J.D., Puvanendran, V., Brown, J.A., 2007. Genetic variation in life-history reaction norms in a marine fish. Proceedings of the Royal Society B: Biological Sciences 274: 1693-1699.
- Hutchinson, W.F., van Oosterhout, C., Rogers, S.I., Carvalho, G.R., 2003. Temporal analysis of archived samples indicates marked genetic changes in declining North Sea cod (*Gadus morhua*). Proceedings of the Royal Society B: Biological Sciences 270: 2125-2132.
- Iles, T.C., Beverton, R.J.H., 2000. The concentration hypothesis: the statistical evidence. ICES Journal of Marine Science: Journal du Conseil 57: 216.
- Jablonski, N.G., 2004. The evolution of human skin and skin color. Annual Review of Anthropology. 33: 585-623.
- Jackson, J.B.C., 2001. Historical overfishing and the recent collapse of coastal ecosystems. Science 293: 629-637.
- 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.
- Jadwiszczak, P., 2009. Rndom Pro 3.14. Software for classical and computer-intensive statistic available free from the New Rndom Site (<http://pjadw.tripod.com>).
- Jakobsdóttir, K.B., Pardoe, H., Magnússon, Á., Björnsson, H., Pampoulie, C., Ruzzante, D.E., Marteinsdóttir, G., 2011. Historical changes in genotypic frequencies at the Pantophysin locus in Atlantic cod (*Gadus morhua*) in Icelandic waters: evidence of fisheries-induced selection? Evolutionary Applications 4: 562-573.



- Jennings, S., Greenstreet, S.P.R., Reynolds J.D., 1999. Structural change in an exploited fish community: a consequence of differential fishing effects on species with contrasting life histories. *Journal of Animal Ecology* 68: 617-627.
- Johnson, D.W., Christie, M.R., Moye, J., Hixon, M.A., 2011. Genetic correlations between adults and larvae in a marine fish: potential effects of fishery selection on population replenishment. *Evolutionary Applications* 4: 621-633.
- Johnson, R.M., Johnson, T.M., Londraville, R.L., 2000. Evidence for leptin expression in fishes. *Journal of Experimental Zoology* 286: 718-724.
- Johnston, S.E., McEwan, J., Pickering, N.K., Kijas, J.W., Beraldi, D., Pilkington, J.G., Pemberton, J.M., Slate, J., 2011. Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*.
- Jonasson, J., Gjerde, B., Gjedrem, T., 1997. Genetic parameters for return rate and body weight of sea-ranched Atlantic salmon. *Aquaculture* 154: 219-231.
- Jones, A.G., Rosenqvist, G., Berglund, A., Avise, J.C., 1999. Clustered microsatellite mutations in the pipefish *Syngnathus typhle*. *Genetics* 152: 1057.
- Jorgensen, C., Enberg, K., Dunlop, E.S., Arlinghaus, R., Boukal, D.S., Brander, K., Ernande, B., Gardmark, A., Johnston, F., Matsumura, S., Pardoe, H., Raab, K., Silva, A., Vainikka, A., Dieckmann, U., Heino, M., Rijnsdorp, A.D., 2007. Managing Evolving Fish Stocks. *Science* 318: 1247-1248.
- Jónsdóttir, I.G., Marteinsdóttir, G., Pampoulie, C., 2008. Relation of growth and condition with the Pan I locus in Atlantic cod (*Gadus morhua* L.) around Iceland. *Mar Biol* 154: 867-874.
- Kadarmideen, H.N., Janss, L.L.G., 2007. Population and systems genetics analyses of cortisol in pigs divergently selected for stress. *Physiological Genomics* 29: 57-65.
- Kalinowski, S.T., Taper, M.L., Marshall, T.C., 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099-1106.
- Kallman, K.D., 1983. The sex determining mechanism of the poeciliid fish, *Xiphophorus montezumae*, and the genetic control of the sexual maturation process and adult size. *Copeia* 1983: 755-769.
- Karim, N., Gordon, S.P., Schwartz, A.K., Hendry, A.P., 2007. This is not déjà vu all over again: male guppy colour in a new experimental introduction. *Journal of Evolutionary Biology* 20: 1339-1350.
- Kauer, M., Dieringer, D., Schlotterer, C., 2003. A microsatellite variability screen for positive selection associated with the "out of Africa" habitat expansion of *Drosophila melanogaster*. *Genetics* 165: 1137.
- Kelly, C.D., Godin, J.G., Wright, J.M., 1999. Geographic variation in multiple paternity within natural populations of the guppy (*Poecilia reticulata*). *Proceedings of the Royal Society of London. Series B: Biological Sciences* 266: 2403.
- Kettlewell, H.B.D., 1955. Selection experiments on industrial melanism in the Lepidoptera. *Heredity* 9: 323-342.
- Kettlewell, H.B.D., 1973. *The Evolution of Melanism*. Clarendon Press, Oxford.

- Khoo, G., Lim, M.H., Suresh, H., Gan, D.K.Y., Lim, K.F., Chen, F., Chan, W.-K., Lim, T.M., Phang, V.P.E., 2003. Genetic linkage maps of the Guppy (*Poecilia reticulata*): assignment of RAPD markers to multipoint linkage groups. *Marine Biotechnology* 5: 279-293.
- Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P., Beerli, P., 2001. The strength of phenotypic selection in natural populations. *The American Naturalist* 157: 245-261.
- Kinnison, M.T., Hendry, A.P., 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* 112: 145-164.
- Kirkpatrick, M., Lande, R., 1989. The evolution of maternal characters. *Evolution* 43: 485-503.
- Korsten, P., Mueller, J.C., Hermannstädter, C., Bouwman, K.M., Dingemanse, N.J., Drent, P.J., Liedvogel, M., Matthysen, E., van Oers, K., van Overveld, T., Patrick, S.C., Quinn, J.L., Sheldon, B.C., Tinbergen, J.M., Kempenaers, B., 2010. Association between DRD4 gene polymorphism and personality variation in great tits: a test across four wild populations. *Molecular Ecology* 19: 832-843.
- Kraak, S.B.M., 2007. Does the probabilistic maturation reaction norm approach disentangle phenotypic plasticity from genetic change? *Marine Ecology Progress Series* 335: 295-300.
- Krohn, M., Reidy, S., Kerr, S., 1997. Bioenergetic analysis of the effects of temperature and prey availability on growth and condition of northern cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences* 54: 113-121.
- Kroon, F.J., Munday, P.L., Westcott, D.A., Hobbs, J.-P.A., Liley, N.R., 2005. Aromatase pathway mediates sex change in each direction. *Proceedings of the Royal Society B: Biological Sciences* 272: 1399-1405.
- Kuparinen, A., Kuikka, S., Merilä, J., 2009. Estimating fisheries-induced selection: traditional gear selectivity research meets fisheries-induced evolution. *Evolutionary Applications* 2: 234-243.
- Kuparinen, A., Merilä, J., 2007. Detecting and managing fisheries-induced evolution. *Trends in Ecology & Evolution* 22: 652-659.
- Kuparinen, A., O'Hara, R.B., Merilä, J., 2008. The role of growth history in determining age and size at maturation in exploited fish populations. *Fish and Fisheries* 9: 201-207.
- Lachance, J., 2009. Detecting selection-induced departures from Hardy-Weinberg proportions. *Genetics Selection Evolution* 41: 15.
- Laland, K.N., Williams, K., 1997. Shoaling generates social learning of foraging information in guppies. *Animal Behaviour* 53: 1161-1169.
- Lampert, K.P., Schmidt, C., Fischer, P., Volf, J.N., Hoffmann, C., Muck, J., Lohse, M.J., Ryan, M.J., Schartl, M., 2010. Determination of onset of sexual maturation and mating behavior by melanocortin receptor 4 polymorphisms. *Current Biology* 20: 1729-1734.
- Lande, R., 1998. Anthropogenic, ecological and genetic factors in extinction and conservation. *Researches on Population Ecology* 40: 259-269.
- Larmuseau, M.H., Raeymakers, J.A., Ruddick, K.G., van Houdt, J.K., Volckart, F.A., 2009. To see in different seas: spatial variation in the rhodopsin gene of the sand goby (*Pomatoschistus minutus*). *Molecular ecology* 18: 4227-4239.

- Larsson, L.C., Laikre, L., Palm, S., André, C., Carvalho, G.R., Ryman, N., 2007. Concordance of allozyme and microsatellite differentiation in a marine fish, but evidence of selection at a microsatellite locus. *Molecular Ecology* 16: 1135–1147.
- Latta, L.C., Bakelar, J.W., Knapp, R.A., Pfrender, M.E., 2007. Rapid evolution in response to introduced predators II: the contribution of adaptive plasticity. *BMC evolutionary biology* 7: 21.
- Law, R., 1979. Life histories under age-specific predation. *The American Naturalist* 114: 399-417.
- Law, R., 2000. Fishing, selection and phenotypic evolution. *ICES Journal of Marine Science* 57: 659-668.
- Law, R., 2007. Fisheries-induced evolution: present status and future directions. Disentangling the causes of maturation trends in exploited fish populations. *Marine Ecology Progress Series* 335: 271–277.
- Law, R., Grey, D.R., 1989. Evolution of yields from populations with age-specific cropping. *Evolutionary Ecology* 3: 343–359.
- Lee, C.E., 2002. Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* 17: 386–391.
- Letcher, B.H., Coombs, J.A., Nislow, K.H., 2011. Maintenance of phenotypic variation: repeatability, heritability and size-dependent processes in a wild brook trout population. *Evolutionary Applications* 4: 602-615.
- Lewontin, R., Cockerham, C.C., 1959. The goodness-of-fit test for detecting natural selection in random mating populations. *Evolution* 13: 561–564.
- Lewontin, R.C., 1974. *The genetic basis of evolutionary change*. Columbia University Press, New York and London.
- Li, M.H., Merilä, J., 2001. Population differences in levels of linkage disequilibrium in the wild. *Molecular Ecology*.
- Liermann, M., Hilborn, R., 2001. Depensation: evidence, models and implications. *Fish and Fisheries* 2: 33–58.
- Lorenzen, K., Enberg, K., 2002. Density-dependent growth as a key mechanism in the regulation of fish populations: evidence from among-population comparisons. *Proceedings of the Royal Society B: Biological Sciences* 269: 49-54.
- Lotze, H.K., Worm, B., 2009. Historical baselines for large marine animals. *Trends in Ecology & Evolution* 24: 254-262.
- Luikart, G., England, P.R., Jordan, S., Taberlet, P., 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics* 4: 981-994.
- Lukas, D., Bradley, B.J., Nsubuga, A.M., Doran-Sheehy, D., Robbins, M.M., Vigilant, L., 2004. Major histocompatibility complex and microsatellite variation in two populations of wild gorillas. *Molecular Ecology* 13: 3389–3402.
- Luttikhuisen, P.C., Drent, J., van Delden, W., Piersma, T., 2003. Spatially structured genetic variation in a broadcast spawning bivalve: quantitative vs. molecular traits. *Journal of Evolutionary Biology* 16: 260–272.
- Lynch, M., Walsh, B., 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.

- Magellan, K., Pettersson, L.B., Magurran, A.E., 2005. Quantifying male attractiveness and mating behaviour through phenotypic size manipulation in the Trinidadian guppy, *Poecilia reticulata*. *Behav Ecol Sociobiol* 58: 366-374.
- Magnhagen, C., 1991. Predation risk as a cost of reproduction. *Trends in Ecology & Evolution* 6: 183-186.
- Magurran, A.E., 2005. *Evolutionary Ecology: The Trinidadian Guppy*, Oxford Series in Ecology and Evolution. Oxford University Press, Oxford.
- Magurran, A.E., Phillip, D.A.T., 2001. Evolutionary implications of large-scale patterns in the ecology of Trinidadian guppies, *Poecilia reticulata*. *Biological Journal of the Linnean Society* 73: 1-9.
- Magurran, A.E., Seghers, B.H., 1991. Variation in schooling and aggression amongst guppy (*Poecilia reticulata*) populations in Trinidad. *Behaviour* 118: 214-234.
- Magurran, A.E., Seghers, B.H., Shaw, P.W., Carvalho, G.R., 1995. The Behavioural Diversity and Evolution of Guppy, *Poecilia reticulata*, Populations in Trinidad. *Advances in the study of behaviour* 24: 155-202.
- Manzon, L.A., 2002. The role of Prolactin in fish osmoregulation: a review. *General and Comparative Endocrinology* 125: 291-310.
- Mardis, E.R., 2008. The impact of next-generation sequencing technology on genetics. *Trends in Genetics* 24: 133-141.
- Marshall, T.C., Slate, J., Kruuk, L.E.B., Pemberton, J.M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* 7: 639-655.
- Martin, N.M., Smith, K.L., Bloom, S.R., Small, C.J., 2006. Interactions between the melanocortin system and the hypothalamo-pituitary-thyroid axis. *Peptides* 27: 333-339.
- Martínez, A., Garza, J.C., Pearse, D.E., 2011. A microsatellite genome screen identifies chromosomal regions under differential selection in Steelhead and Rainbow Trout. *Transactions of the American Fisheries Society* 140: 829-842.
- Matthews, I.M., Evans, J.P., Magurran, A.E., 1997. Male display rate reveals ejaculate characteristics in the Trinidadian guppy, *Poecilia reticulata*. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 264: 695.
- Mattingly, H.T., Butler, M.J., 1994. Laboratory predation on the Trinidadian guppy: implications for the size-selective predation hypothesis and guppy life history evolution. *Oikos* 69: 54-64.
- Mayr, E., 1966. *Animal species and evolution*. Harvard University Press, Cambridge, MA.
- McKay, J.K., Latta, R.G., 2002. Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution* 17: 285-291.
- McKenzie, J.A., Batterham, P., 1994. The genetic, molecular and phenotypic consequences of selection for insecticide resistance. *Trends in Ecology & Evolution* 9: 166-169.
- McPherron, A., Lawler, A.M., Lee, S.-J., 1997. Regulation of skeletal muscle mass in mice by a new TGF-beta superfamily member. *Nature* 387: 83-90.
- Meirmans, P.G., Hedrick, P.W., 2011. Assessing population structure: FST and related measures. *Molecular Ecology Resources*.

- Meirmans, P.G., van Tienderen, P.H., 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4: 792–794.
- Merilä, J., Crnokrak, P., 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology* 14: 892–903.
- Merilä, J., Kruuk, L.E.B., Sheldon, B.C., 2001a. Cryptic evolution in a wild bird population. *Nature* 412: 76–79.
- Merilä, J., Sheldon, B.C., Kruuk, L.E.B., 2001b. Explaining stasis: microevolutionary studies in natural populations. *Genetica* 112: 199–222.
- Mertz, G., Myers, R.A., 1998. A simplified formulation for fish production. *Canadian Journal of Fisheries and Aquatic Sciences* 55: 478–484.
- Miller, R.M., 1957. Have the genetic patterns of fishes been altered by introductions or by selective fishing? *Journal of the Fisheries Research Board of Canada* 14: 797–806.
- Miyata, A., Arimura, A., Dahl, R.R., Minamino, N., Uehara, A., Jiang, L., et al., 1989. Isolation of a novel 38 residue-hypothalamic polypeptide which stimulates adenylate cyclase in pituitary cells. *Biochemical and Biophysical Research Communications* 164: 567–574.
- Mollet, F.M., Kraak, S.B.M., Rijnsdorp, A.D., 2006. Fisheries-induced evolutionary changes in maturation reaction norms in North Sea sole (*Solea solea*).
- Montgomery, M.E., Woodworth, L.M., Nurthen, R.K., Gilligan, D.M., Briscoe, D.A., Frankham, R., 2000. Relationships between population size and loss of genetic diversity: comparisons of experimental results with theoretical predictions. *Conservation Genetics* 1: 33–43.
- Morgan, T.J., Garland Jr., T., Irwin, B.L., Swallow, J.G., Carter, P.A., 2003. The mode of evolution of molecular markers in populations of house mice under artificial selection for locomotor behavior. *Journal of Heredity* 94: 236–242.
- Morita, K., Fukawaka, M., 2006. Does size matter most? The effect of growth history on probabilistic reaction norm for salmon maturation. *Evolution* 60: 1516–1521.
- Morita, K., Tsuboi, J.-i., Nagasawa, T., 2009. Plasticity in probabilistic reaction norms for maturation in a salmonid fish. *Biology Letters* 5: 628–631.
- Moriyama, S., Ayson, F.G., Kawauchi, H., 2000. Growth regulation by Insulin-like Growth Factor-I in fish. *Bioscience, Biotechnology and Biochemistry* 64: 1553–1562.
- Munch, S.B., Walsh, M.R., Conover, D.O., 2005. Harvest selection, genetic correlations, and evolutionary changes in recruitment: one less thing to worry about? *Canadian Journal of Fisheries and Aquatic Sciences* 62: 802–810.
- Myers, R.A., Barrowman, N.J., Hutchings, J.A., Rosenberg, A.A., 1995. Population dynamics of exploited fish stocks at low population levels. *Science* 269: 1106–1108.
- Myers, R.A., Hutchings, J.A., Barrowman, N.J., 1997. Why do fish stocks collapse? The example of Cod in Atlantic Canada. *Ecological Applications* 7: 91–106.
- Myers, R.A., Worm, B., 2003. Rapid worldwide depletion of predatory fish communities. *Nature* 423: 280–283.
- Møller, A.P., Mousseau, T.A., 2008. Reduced abundance of raptors in radioactively contaminated areas near Chernobyl. *Journal of Ornithology* 150: 239–246.
- Møller, A.P., Mousseau, T.A., 2009. Reduced abundance of insects and spiders linked to radiation at Chernobyl 20 years after the accident. *Biology Letters* 5: 356–359.

- Neff, B.D., Pitcher, T.E., Ramnarine, I.W., 2008. Inter-population variation in multiple paternity and reproductive skew in the guppy. *Molecular Ecology* 17: 2975-2984.
- Nei, M., 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York, USA.
- Nielsen, E., Hemmer-Hansen, J., Poulsen, N., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G.L., Ogden, R., Carvalho, G., 2009b. Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology* 9: 276.
- Nielsen, E.E., Hansen, M.M., Bach, L.A., 2001. Looking for a needle in a haystack: discovery of indigenous Atlantic salmon (*Salmo salar* L.) in stocked populations. *Conservation Genetics* 2: 219-232.
- Nielsen, E.E., Hansen, M.M., Meldrup, D., 2006. Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. *Molecular Ecology* 15: 3219-3229.
- Nielsen, E.E., Hemmer-Hansen, J., Larsen, P.F., Bekkevold, D., 2009a. Population genomics of marine fishes: identifying adaptive variation in space and time. *Molecular Ecology* 18: 3128-3150.
- Nusslé, S., Bornand, C.N., Wedekind, C., 2009. Fishery-induced selection on an Alpine whitefish: quantifying genetic and environmental effects on individual growth rate. *Evolutionary Applications* 2: 200-208.
- Nusslé, S., Bréchon, A., Wedekind, C., 2011. Change in individual growth rate and its link to gill-net fishing in two sympatric whitefish species. *Evolutionary Ecology* 25: 681-693.
- Olendorf, R., Reudi, B., Hughes, K.A., 2004. Primers for 12 polymorphic microsatellite DNA loci from the guppy (*Poecilia reticulata*). *Molecular Ecology Notes* 4: 668-671.
- Oliver, M.K., Telfer, S., Piertney, S.B., 2009. Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water vole (*Arvicola terrestris*). *Proceedings of the Royal Society B: Biological Sciences* 276: 1119-1128.
- Olsen, E.M., Heino, M., Lilly, G.R., Morgan, M.J., Brattey, J., Ernande, B., Dieckmann, U., 2004. Maturation Trends indicative of rapid evolution preceded the collapse of northern cod. *Nature* 428: 932-935.
- Ono, M., Takayama, Y., Rand-Weaver, M., Sakata, S., Yasunaga, T., Noso, T., Kawauchi, H., 1990. cDNA cloning of somatolactin, a pituitary protein related to growth hormone and prolactin. *Proceedings of the National Academy of Sciences* 87: 4330-4334.
- van Oosterhout, C., Harris, P.D., Cable, J., 2003. Marked variation in parasite resistance between two wild populations of the Trinidadian guppy, *Poecilia reticulata* (Pisces: Poeciliidae). *Biological Journal of the Linnean Society* 79: 645-651.
- van Oosterhout, C., Hutchinson, W.F., Wills, D.P., Shipley, P., 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535-538.
- van Oosterhout, C., Joyce, D.A., Cummings, S.M., 2006b. Evolution of MHC class IIB in the genome of wild and ornamental guppies, *Poecilia reticulata*. *Heredity* 97: 111-118.
- van Oosterhout, C., Joyce, D.A., Cummings, S.M., Blais, J., Barson, N.J., Ramnarine, I.W., Mohammed, R.S., Persad, N., Cable, J., 2006a. Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (*Poecilia reticulata*). *Evolution* 60: 2562-2574.

- van Oosterhout, C., Mohammed, R.S., Hansen, H., Archard, G.A., McMullan, M., Weese, D.J., Cable, J., 2007b. Selection by parasites in spate conditions in wild Trinidadian guppies (*Poecilia reticulata*). *International Journal for Parasitology* 37: 805-812.
- van Oosterhout, C., Smith, A.M., Hänfling, B., Ramnarine, I.W., Mohammed, R.S., Cable, J., 2007a. The guppy as a conservation model: implications of parasitism and inbreeding for reintroduction success. *Conservation Biology* 21: 1573-1583.
- van Oosterhout, C., Trigg, R.E., Carvalho, G.R., Magurran, A.E., Hauser, L., Shaw, P.W., 2003. Inbreeding depression and genetic load of sexually selected traits: how the guppy lost its spots. *Journal of Evolutionary Biology* 16: 273–281.
- Palkovacs, E.P., Dion, K.B., Post, D.M., Caccone, A., 2008. Independent evolutionary origins of landlocked alewife populations and rapid parallel evolution of phenotypic traits. *Molecular Ecology* 17: 582–597.
- Palkovacs, E.P., Wasserman, B.A., Kinnison, M.T., 2011. Eco-evolutionary trophic dynamics: loss of top predators drives trophic evolution and ecology of prey. *PLoS ONE* 6: e18879.
- Palumbi, S.R., 2001. Humans as the World’s Greatest Evolutionary Force. *Science* 293: 1786-1790.
- Park, L.K., Moran, P., 1994. Developments in molecular genetic techniques in fisheries. *Reviews in Fish Biology and Fisheries* 4: 272–299.
- Parsell, D., Lindquist, S., 1993. The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. *Annual Review of Genetics* 27: 437–496.
- Paterson, I.G., Crispo, E., Kinnison, M.T., Hendry, A.P., Bentzen, P., 2005. Characterization of tetranucleotide microsatellite markers in guppy (*Poecilia reticulata*). *Molecular Ecology Notes* 5: 269–271.
- Pauly, D., 1995. Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology and Evolution* 10: 430.
- Pauly, D., Alder, J., Bennett, E., Christensen, V., Tyedmers, P., Watson, R., 2003. The future for fisheries. *Science* 302: 1359.
- Peters, W.C.H., 1859. Eine neue vom herrn jagor im Atlantischen Meere gefangene art der gattung *Leptocephalus*, und über einige andere neue fisch des Zoologischen Museums. *Monatsbericht der Konigliche Akademie der Wissenschaften zu Berlin* 411-413.
- Petitgas, P., Secor, D.H., McQuinn, I., Huse, G., Lo, N., 2010. Stock collapses and their recovery: mechanisms that establish and maintain life-cycle closure in space and time. *ICES Journal of Marine Science: Journal du Conseil* 67: 1841.
- Piertney, S.B., Webster, L.M.I., 2010. Characterising functionally important and ecologically meaningful genetic diversity using a candidate gene approach. *Genetica* 138: 419-432.
- Pinsky, M.L., Jensen, O.P., Ricard, D., Palumbi, S.R., 2011. Unexpected patterns of fisheries collapse in the world’s oceans. *Proceedings of the National Academy of Sciences*.
- Pogson, G.H., 2001. Nucleotide polymorphism and natural selection at the Pantophysin (Pan I) locus in the Atlantic cod, *Gadus morhua* (L.). *Genetics* 157: 317-330.
- Policansky, D., 1993. Fishing as a cause of evolution in fishes. In: *The Exploitation of Evolving Resources*. Springer-Verlag, Berlin, pp. 2-18.

- Poulsen, N.A., Nielsen, E.E., Schierup, M.H., Loeschcke, V., Gronkjaer, P., 2006. Long-term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*). *Molecular Ecology* 15: 321-331.
- Primmer, C.R., Borge, T., Lindell, J., Saetre, G.P., 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Molecular Ecology* 11: 603-612.
- Rand, D.M., Weinreich, D.M., Lerman, D., Folk, D., Gilchrist, G.W., 2010. Three selections are better than one: Clinal variation of thermal QTL from independent selection experiments in *Drosophila*. *Evolution* 64: 2921-2934.
- Rand-Weaver, M., Swanson, P., Kawachi, H., Dickhoff, W.W., 1992. Somatolactin, a novel pituitary protein: purification and plasma levels during reproductive maturation of coho salmon. *Journal of Endocrinology* 133: 393-403.
- Raymond, M., Berticat, C., Weill, M., Pasteur, N., Chevillon, C., 2001. Insecticide resistance in the mosquito *Culex pipiens*: what have we learned about adaptation? *Genetica* 112: 287-296.
- Raymond, M., Rousset, F., 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248-249.
- Reis, E.G., Pawson, M.G., 1999. Fish morphology and estimating selectivity by gillnets. *Fisheries research* 39: 263-273.
- Renaut, S., Nolte, A.W., Rogers, S.M., Derome, N., Bernatchez, L., 2011. SNP signatures of selection on standing genetic variation and their association with adaptive phenotypes along gradients of ecological speciation in lake whitefish species pairs (*Coregonus* spp.). *Molecular Ecology* 20: 545-559.
- Reusch, T.B., Wood, T.E., 2007. Molecular ecology of global change. *Molecular Ecology* 16: 3973-3992.
- Reynolds, J.D., Gross, M.R., 1992. Female mate preference enhances offspring growth and reproduction in a fish, *Poecilia reticulata*. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 250: 57-62.
- Reznick, D.N., 1982. The impact of predation on life history evolution in Trinidadian guppies: genetic basis of observed life history patterns. *Evolution* 36: 1236-1250.
- Reznick, D.N., 1989. Life-history evolution in guppies: 2. Repeatability of field observations and the effects of season on life histories. *Evolution* 43: 1285-1297.
- Reznick, D.N., 1990. Plasticity in age and size at maturity in male guppies (*Poecilia reticulata*): An experimental evaluation of alternative models of development. *Journal of Evolutionary Biology* 3: 185-203.
- Reznick, D.N., 1997. Life history evolution in guppies (*Poecilia reticulata*): Guppies as a model for studying the evolutionary biology of aging. *Experimental Gerontology* 32: 245-258.
- Reznick, D.N., Bryant, M.J., Roff, D.A., Ghalambor, C.K., Ghalambor, D.E., 2004. Effect of extrinsic mortality on the evolution of senescence in guppies. *Nature* 431: 1095-1099.
- Reznick, D.N., Bryga, H.A., 1987. Life-history evolution in guppies (*Poecilia reticulata*): 1. Phenotypic and genetic changes in an introduction experiment. *Evolution* 41: 1370-1385.
- Reznick, D.N., Bryga, H.A., 1996. Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). V. Genetic basis of parallelism in life histories. *The American Naturalist* 147: 339-359.



- Reznick, D.N., Bryga, H.A., Endler, J.A., 1990. Experimentally induced life-history evolution in a natural population. *Nature* 346: 357-359.
- Reznick, D.N., Buckwalter, G., Groff, J., Elder, D., 2001b. The evolution of senescence in natural populations of guppies (*Poecilia reticulata*): a comparative approach. *Experimental Gerontology* 36: 791-812.
- Reznick, D.N., Butler, M.J., Rodd, F.H., Ross, P., 1996. Life-history evolution in guppies (*Poecilia reticulata*) 6. Differential mortality as a mechanism for natural selection. *Evolution* 50: 1651-1660.
- Reznick, D.N., Butler, M.J., Rodd, H., 2001a. Life-history evolution in guppies. VII. The comparative ecology of high- and low-predation environments. *The American Naturalist* 157: 126-140.
- Reznick, D.N., Endler, J.A., 1982. The impact of predation on life history evolution in Trinidadian Guppies (*Poecilia reticulata*). *Evolution* 36: 160-177.
- Reznick, D.N., Ghalambor, C.K., 2001. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* 112: 183-198.
- Reznick, D.N., Ghalambor, C.K., 2005a. Can commercial fishing cause evolution? Answers from guppies (*Poecilia reticulata*). *Canadian Journal of Fisheries and Aquatic Sciences* 62: 791-801.
- Reznick, D.N., Ghalambor, C.K., 2005b. Selection in nature: experimental manipulations of natural populations. *Integrative and Comparative Biology* 45: 456.
- Reznick, D.N., Miles, D.B., Winslow, S., 1992. Life history of *Poecilia picta* (Poeciliidae) from the island of Trinidad. *Copeia* 1992: 782-790.
- Reznick, D.N., Shaw, F.H., Rodd, F.H., Shaw, R.G., 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* 275: 1934-1937.
- Ricker, W.E., 1981. Changes in the average size and average age of Pacific Salmon. *Canadian Journal of Fisheries and Aquatic Sciences* 38: 1636-1656.
- Ricker, W.E., Bilton, H.T., Aro, K.V., 1978. Causes of the decrease in size of pink salmon (*Oncorhynchus gorbuscha*). Fisheries and Marine Service Technical Report (No. 820). Department of Fisheries and Environment, Canada.
- Rijnsdorp, A.D., 1993. Fisheries as a large-scale experiment on life-history evolution: disentangling phenotypic and genetic effects in changes in maturation and reproduction of North Sea plaice, *Pleuronectes platessa* L. *Oecologia* 96: 391-401.
- Rijnsdorp, A.D., Daan, N., Dekker, W., Poos, J., Vandensen, W., 2007. Sustainable use of flatfish resources: Addressing the credibility crisis in mixed fisheries management. *Journal of Sea Research* 57: 114-125.
- Rochet, M.J., 1998. Short-term effects of fishing on life history traits of fishes. *ICES Journal of Marine Science: Journal du Conseil* 55: 371.
- Rodd, F.H., Reznick, D.N., 1991. Life history evolution in guppies: III. The impact of prawn predation on guppy life histories. *Oikos* 62: 13-19.
- Rodd, F.H., Reznick, D.N., 1997. Variation in the demography of guppy populations: the importance of predation and life histories. *Ecology Letters* 78: 405-418.
- Roff, D.A., 1991. The evolution of life-history variation in fishes, with particular reference to flatfishes. *Netherlands Journal of Sea Research* 27: 197-207.

- Roff, D.A., 1992. *The Evolution of Life Histories: Theory and Analysis*. Chapman and Hall, New York.
- Rogers, S.M., Bernatchez, L., 2005. Integrating QTL mapping and genome scans towards the characterization of candidate loci under parallel selection in the lake whitefish (*Coregonus clupeaformis*). *Molecular Ecology* 14: 351-361.
- de Roos, A.M., Boukal, D.S., Persson, L., 2006. Evolutionary regime shifts in age and size at maturation of exploited fish stocks. *Proceedings of the Royal Society B: Biological Sciences* 273: 1873-1880.
- Rose, K.A., Cowan, J.H., Winemiller, K.O., Myers, R.A., Hilborn, R., 2001. Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis. *Fish and Fisheries* 2: 293-327.
- Rose, M.R., Passananti, H.B., Chippindale, A.K., Phelan, J.P., Matos, M., Teotonio, T., Mueller, M.D., 2005. The effects of evolution are local: Evidence from experimental evolution in *Drosophila*. *Integrative and Comparative Biology* 45: 486-491.
- Rousset, F., 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* 8: 103-106.
- Le Rouzic, A.L., Ostbye, K., Klepaker, T.O., Hansen, T.F., Bernatchez, L., Schluter, D., Völlestad, L.A., 2011. Strong and consistent natural selection associated with armour reduction in sticklebacks. *Molecular Ecology*.
- Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology* 132: 365. Available at: [http://biotoools.umassmed.edu/bioapps/primer3\\_www.cgi](http://biotoools.umassmed.edu/bioapps/primer3_www.cgi)
- Russello, M.A., Kirk, S.L., Frazer, K.A., Askey, P., 2011. Detection of outlier loci and their utility for fisheries management. *Evolutionary Applications*. In press. DOI: 10.1111/j.1752-4571.2011.00206.x
- Rutter, C., 1902. Natural history of the Quinnet salmon. A report on investigations in the Sacramento River. *Bulletin of the U.S. Fish Commission* 22: 1886-1901.
- Ruzzante, D.E., Doyle, R.W., 1991. Rapid behavioural changes in Medaka (*Oryzias latipes*) caused by selection for competitive and noncompetitive growth. *Evolution* 45: 1936-1946.
- Ruzzante, D.E., Mariani, S., Bekkevold, D., Andre, C., Mosegaard, H., Clausen, L.A., Dahlgren, T.G., Hutchinson, W.F., Hatfield, E.M., Torstensen, E., Brigham, J., Simmonds, E.J., Laikre, L., Larsson, L.C., Stet, R.J., Ryman, N., Carvalho, G.R., 2006. Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring. *Proceedings of the Royal Society B: Biological Sciences* 273: 1459-1464.
- Ruzzante, D.E., Wroblewski, J.S., Taggart, C.T., Smedbol, R.K., Cook, D., Goddaard, S.V., 2000. Bay-scale population structure in coastal Atlantic cod in Labrador and Newfoundland, Canada. *Journal of Fish Biology* 56: 431-447.
- Ryman, N., Palm, S., André, C., Carvalho, G.R., Dahlgren, T.G., Jorde, P.E., Laikre, L., Larsson, L.C., Palmé, A., Ruzzante, D.E., 2006. Power for detecting genetic divergence: differences between statistical methods and marker loci. *Molecular Ecology* 15: 2031-2045.
- Ryynanen, H.J., Primmer, C.R., 2004. Primers for sequence characterization and polymorphism detection in the Atlantic salmon (*Salmo salar*) growth hormone 1 (GH1) gene. *Molecular Ecology Notes* 4: 664-667.

- Salem, M., Silverstein, J., Rexroad, C., Yao, J., 2007. Effect of starvation on global gene expression and proteolysis in rainbow trout (*Oncorhynchus mykiss*). *BMC genomics* 8: 328.
- Sato, A., Figueroa, F., O'hUigin, C., Reznick, D.N., Klein, J., 1995. Identification of major histocompatibility complex genes in the guppy, *Poecilia reticulata*. *Immunogenetics* 43: 38-49.
- Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., Walker, B., others, 2001. Catastrophic shifts in ecosystems. *Nature* 413: 591–596.
- Scheiner, S.M., 2002. Selection experiments and the study of phenotypic plasticity. *Journal of Evolutionary Biology* 15: 889–898.
- Schlingensiepen, K.H., Wollnik, F., Kunst, M., Schlingensiepen, R., Herdegen, T., Brysch, W., 1994. The role of Jun transcription factor expression and phosphorylation in neuronal differentiation, neuronal cell death, and plastic adaptations in vivo. *Cellular and molecular neurobiology* 14: 487–505.
- Schlötterer, C., 2002. A microsatellite-based multilocus screen for the identification of local selective sweeps. *Genetics* 160: 753.
- Schlötterer, C., 2003. Hitchhiking mapping-functional genomics from the population genetics perspective. *Trends in Genetics* 19: 32–38.
- Schmidt, J., 1919. Racial studies in fishes. *Journal of genetics* 8: 147–153.
- Schories, S., Meyer, M.K., Schartl, M., 2009. Description of *Poecilia (Acanthophaelus) obscura* n. sp., (Teleostei: Poeciliidae), a new guppy species from western Trinidad, with remarks on *P. wingei* and the status of the “Endler’s guppy”. *Zootaxa* 2266: 35-50.
- Schuelke, M., 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.
- Schwartz, A.K., Hendry, A.P., 2010. Testing the influence of local forest canopy clearing on phenotypic variation in Trinidadian guppies. *Functional Ecology* 24: 354–364.
- Shackell, N.L., Frank, K.T., Fisher, J.A.D., Petrie, B., Leggett, W.C., 2009. Decline in top predator body size and changing climate alter trophic structure in an oceanic ecosystem. *Proceedings of the Royal Society B: Biological Sciences* 277: 1353-1360.
- Sharpe, D.M., Hendry, A.P., 2009. Life history change in commercially exploited fish stocks: an analysis of trends across studies. *Evolutionary Applications* 2: 260–275.
- Shaw, P.W., Carvalho, G.R., Magurran, A.E., Seghers, B.H., 1991. Population differentiation in Trinidadian guppies (*Poecilia reticulata*): patterns and problems. *Journal of Fish Biology* 39: 203–209.
- Shaw, P.W., Carvalho, G.R., Magurran, A.E., Seghers, B.H., 1994. Factors affecting the distribution of genetic variability in the guppy, *Poecilia reticulata*. *Journal of Fish Biology* 45: 875-888.
- Shen, X., Yang, G., Liao, M., 2007a. Development of 51 genomic microsatellite DNA markers of guppy (*Poecilia reticulata*) and their application in closely related species. *Molecular Ecology Notes* 7: 302–306.
- Shen, X., Yang, G., Liu, Y., Liao, M., Wang, X., Zhu, M., Song, W., Zou, G., Wei, Q., Wang, D., Chen, D., 2007b. Construction of genetic linkage maps of guppy (*Poecilia reticulata*) based on AFLP and microsatellite DNA markers. *Aquaculture* 271: 178–187.

- Silliman, 1975: 1975. Selective and unselective exploitation of experimental populations of *Tilapia mossambica*. Fishery Bulletin of the US 73: 495-507.
- Simões, P., Pascual, M., Coelho, M.M., Matos, M., 2010. Divergent evolution of molecular markers during laboratory adaptation in *Drosophila subobscura*. Genetica 138: 999-1009.
- Simões, P., Pascual, M., Santos, J., Rose, M., Matos, M., 2008a. Evolutionary dynamics of molecular markers during local adaptation: a case study in *Drosophila subobscura*. BMC evolutionary biology 8: 66.
- Simões, P., Pascual, M., Santos, J., Rose, M., Matos, M., 2008b. Correction: Evolutionary dynamics of molecular markers during local adaptation: a case study in *Drosophila subobscura*. BMC evolutionary biology 8: 66.
- Slate, J., Gratten, J., Beraldi, D., Stapley, J., Hale, M., Pemberton, J.M., 2008. Gene mapping in the wild with SNPs: guidelines and future directions. Genetica 136: 97-107.
- Smith, P.J., Francis, R., McVeagh, M., 1991. Loss of genetic diversity due to fishing pressure. Fisheries Research 10: 309–316.
- Smoker, W.W., Gharrett, A.J., Stekoll, M.S., Taylor, S.G., 2000. Genetic variation of fecundity and egg size in anadromous pink salmon, *Oncorhynchus gorbuscha* Walbaum. Alaska Fisheries Research Bulletin 7: 44-50.
- Solmundsson, J., Karlsson, H., Palsson, J., 2003. Sexual differences in spawning behaviour and catchability of plaice (*Pleuronectes platessa*) west of Iceland. Fisheries Research 61: 57–71.
- Solé, X., Guinó, E., Valls, J., Iniesta, R., Moreno, V., 2006. SNPStats: a web tool for the analysis of association studies. Bioinformatics 22: 1928. Available at: [http://bioinfo.iconcologia.net/SNPstats\\_web](http://bioinfo.iconcologia.net/SNPstats_web)
- St Juliana, J.R., Janzen, F.J., 2007. Can natural phenotypic variances be estimated reliably under homogeneous laboratory conditions? Journal of Evolutionary Biology 20: 1406-1414.
- Stapley, J., Reger, J., Feulner, P.G.D., Smadja, C., Galindo, J., Ekblom, R., Bennison, C., Ball, A.D., Beckerman, A.P., Slate, J., 2010. Adaptation genomics: the next generation. Trends in Ecology & Evolution 25: 705-712.
- Stearns, S.C., 1983. A natural experiment in life-history evolution: field data on the introduction of mosquitofish (*Gambusia affinis*) to Hawaii. Evolution 601–617.
- Stearns, S.C., 1989. The evolutionary significance of phenotypic plasticity. BioScience 39: 436–445.
- Stearns, S.C., Koella, J.C., 1986. The evolution of phenotypic plasticity in life-history traits: predictions of reaction norms for age and size at maturity. Evolution 40: 893–913.
- Steingrund, P., Gaard, E., 2005. Relationship between phytoplankton production and cod production on the Faroe shelf. ICES Journal of Marine Science 62: 163-176.
- Stenseth, N.C., Dunlop, E.S., 2009. Unnatural selection. Nature 457: 803–804.
- Stinchcombe, J.R., Hoekstra, H.E., 2007. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity 100: 158–170.
- Stockwell, C.A., Hendry, A.P., Kinnison, M.T., 2003. Contemporary evolution meets conservation biology. Trends in Ecology & Evolution 18: 94-101.

- Stokes, K., Law, R., 2000. Fishing as an evolutionary force. *Marine Ecology Progress Series* 208: 307-309.
- Stokes, T.K., McGlade, J.M., Law, R., 1993. *The Exploitation of Evolving Resources*. Springer-Verlag, Berlin.
- Storz, J.F., 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology* 14: 671–688.
- Suk, H.Y., Neff, B.D., 2009. Microsatellite genetic differentiation among populations of the Trinidadian guppy. *Heredity* 102: 425–434.
- Svetec, N., Werzner, A., Wilches, R., Pavlidis, P., Alvarez-Castro, J., Broman, K.W., Metzler, D., Stephan, W., 2011. Identification of X-linked quantitative trait loci affecting cold tolerance in *Drosophila melanogaster* and fine mapping by selective sweep analysis. *Molecular Ecology* 20: 530-544.
- Swain, D.P., 1993. Age- and density-dependent bathymetric pattern of Atlantic cod (*Gadus morhua*) in the southern Gulf of St Lawrence. *Canadian Journal of Fisheries and Aquatic Sciences* 50: 1255-1264.
- Swain, D.P., 2011. Life-history evolution and elevated natural mortality in a population of Atlantic cod (*Gadus morhua*). *Evolutionary Applications* 4: 18-29.
- Swain, D.P., Sinclair, A.F., Hanson, M.J., 2007. Evolutionary response to size-selective mortality in an exploited fish population. *Proceedings of the Royal Society B: Biological Sciences* 274: 1015-1022.
- Swallow, J.G., Carter, P.A., Garland, T., 1998. Artificial selection for increased wheel-running behavior in house mice. *Behavior genetics* 28: 227–237.
- Sørensen, J.G., Nielsen, M.M., Loeschcke, V., 2007. Gene expression profile analysis of *Drosophila melanogaster* selected for resistance to environmental stressors. *J Evolution Biol* 20: 1624-1636.
- Tao, W.J., Boulding, E.G., 2003. Associations between single nucleotide polymorphisms in candidate genes and growth rate in Arctic charr (*Salvelinus alpinus* L.). *Heredity* 91: 60–69.
- Tenaillon, M., Tiffin, P., 2008. The quest for adaptive evolution: a theoretical challenge in a maze of data. *Current Opinion in Plant Biology* 11: 110-115.
- Tonteri, A., Vasemägi, A., Lumme, J., Primmer, C.R., 2010. Beyond MHC: signals of elevated selection pressure on Atlantic salmon (*Salmo salar*) immune-relevant loci. *Molecular Ecology* 19: 1273–1282.
- Tripathi, N., Hoffmann, M., Weigel, D., Dreyer, C., 2009b. Linkage analysis reveals the independent origin of Poeciliid sex chromosomes and a case of atypical sex inheritance in the Guppy (*Poecilia reticulata*). *Genetics* 182: 365-374.
- Tripathi, N., Hoffmann, M., Willing, E.M., Lanz, C., Weigel, D., Dreyer, C., 2009a. Genetic linkage map of the guppy, *Poecilia reticulata*, and quantitative trait loci analysis of male size and colour variation. *Proceedings of the Royal Society B: Biological Sciences* 276: 2195-2208.
- Trippel, E.A., 1995. Age at maturity as a stress indicator in fisheries. *BioScience* 45: 759–771.
- Tuggle, C.K., Trenkle, A., 1996. Control of growth hormone synthesis. *Domestic Animal Endocrinology* 13: 1-33.
- Tysklind, N., 2009. Population genetic markers in biomonitoring programmes: a case study of flatfish around the British Isles. (PhD Thesis).

- Unniappan, S., Peter, R.E., 2005. Structure, distribution and physiological functions of ghrelin in fish. *Comparative Biochemistry and Physiology-Part A: Molecular & Integrative Physiology* 140: 396–408.
- Uusi-Heikkilä, S., Wolter, C., Klefoth, T., Arlinghaus, R., 2008. A behavioral perspective on fishing-induced evolution. *Trends in Ecology & Evolution* 23: 419–421.
- Vasemägi, A., Primmer, C.R., 2005. Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology* 14: 3623–3642.
- Vitalis, R., Dawson, K., Boursot, P., 2001. Interpretation of variation across marker loci as evidence of selection. *Genetics* 158: 1811–1823.
- Volff, J.N., Schartl, M., 2001. Variability of genetic sex determination in Poeciliid fishes. *Genetica* 111: 101–110.
- Walsh, M.R., Munch, S.B., Chiba, S., Conover, D.O., 2006. Maladaptive changes in multiple traits caused by fishing: impediments to population recovery. *Ecology Letters* 9: 142–148.
- Watanabe, T., Nakajima, M., Yoshida, M., Taniguchi, N., 2004. Construction of six linkage groups in the guppy (*Poecilia reticulata*). *Animal Genetics* 35: 142–167.
- Watanabe, T., Nakajima, Y.M., Taniguchi, N., 2003. Isolation and characterization of 43 microsatellite DNA markers for guppy (*Poecilia reticulata*). *Molecular Ecology Notes* 3: 487–490.
- Watson, C.T., Gray, S.M., Hoffmann, M., Lubieniecki, K.P., Joy, J.B., Sandkam, B.A., Weigel, D., Loew, E., Dreyer, C., Davidson, W.S., Breden, F., 2011. Gene duplication and divergence of long wavelength-sensitive opsin genes in the Guppy, *Poecilia reticulata*. *J Molecular Evolution* 72: 240–252.
- Watson, R., Pauly, D., 2001. Systematic distortions in world fisheries catch trends. *Nature* 414: 534–536.
- Watterson, G., 1977. Heterosis or neutrality? *Genetics* 85: 789.
- Watterson, G., 1978. The homozygosity test of neutrality. *Genetics* 88: 405.
- Weigensberg, I., Roff, D.A., 1996. Natural heritabilities: can they be reliably estimated in the laboratory? *Evolution* 50: 2149–2157.
- Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
- Wells, H.G., Huxley, J., Wells, G.P., 1934. *Evolution - Fact and Theory*. Cassell, London, UK.
- Wheat, C.W., Fescemeyer, H.W., Kvist, J., Tas, E., Vera, J.C., Frilander, M.J., Hanski, I., Marden, J.H., 2011. Functional genomics of life history variation in a butterfly metapopulation. *Molecular Ecology*.
- Willing, E.M., Bentzen, P., van Oosterhout, C., Hoffmann, M., Cable, J., Breden, F., Weigel, D., Dreyer, C., 2010. Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. *Molecular ecology* 19: 968–984.
- Winemiller, K.O., Rose, K.A., 1992. Patterns of life-history diversification in North American fishes: Implications for population regulation. *Canadian Journal of Fisheries and Aquatic Sciences* 49: 2196–2218.
- Winge, Ö., 1922a. A peculiar mode of inheritance and its cytological explanation. *Journal of Genetics* 12: 137–144.

- Winge, Ö., 1922b. One-sided masculine and sex-linked inheritance in *Lebistes reticulatus*. *Journal of Genetics* 12: 145–162.
- Winge, Ö., 1937. Succession of Broods in *Lebistes*. *Nature* 140: 467.
- Witkovsky, P., 2004. Dopamine and retinal function. *Documenta Ophthalmologica* 108: 17-39.
- Wood, H.M., Grahame, J.W., Humphray, S., Rogers, J., Butlin, R.K., 2008. Sequence differentiation in regions identified by a genome scan for local adaptation. *Molecular Ecology* 17: 3123–3135.
- Worm, B., Barbier, E., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S., Jackson, J.B.C., Lotze, H.K., Micheli, F., Palumbi, S.R., Sala, E., Selkoe, K.A., Stachowicz, J.J., Watson, R., 2006. Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787-790.
- Wright, S., 1951. The genetical structure of natural populations. *Annals of Eugenics* 15: 323-354.
- Wright, S., 1977. Experimental results and evolutionary deductions, evolution and the genetics of populations. University of Chicago Press, Chicago, USA.
- Xiong, M., Jin, L., 1999. Comparison of the power and accuracy of biallelic and microsatellite markers in population-based gene-mapping methods. *The American Journal of Human Genetics* 64: 629–640.
- Xu, Y.X., Zhu, Z.Y., Lo, L.C., Wang, C.M., Lin, G., Feng, F., Yue, G.H., 2006. Characterization of two parvalbumin genes and their association with growth traits in Asian seabass (*Lates calcarifer*). *Animal Genetics* 37: 266-268.
- Young, C.W., Legates, J.E., Farthing, B.R., 1965. Prenatal and postnatal influences on growth, prolificacy and maternal performance in mice. *Genetics* 52: 553-561.
- Zhou, S., Smith, A.D.M., Punt, A.E., Richardson, A.J., Gibbs, M., Fulton, E.A., Pascoe, S., Bulman, C., Bayliss, P., Sainsbury, K., 2010. Ecosystem-based fisheries management requires a change to the selective fishing philosophy. *Proceedings of the National Academy of Sciences* 107: 9485-9489.

## **APPENDICES**



## **APPENDIX I – Primers and sequences**

- a. Microsatellite primer pairs showing marker name, annealing temperature used (AT), forward (F) and reverse (R) primer sequence, including a universal tail on each forward primer, complementary to a fluorescently labelled universal primer. The colour of the tail is shown under F tail, and its sequence shown at the bottom of the table, as well as in small letters included in each respective forward primer.
- b. Candidate primer sequences obtained from C. Dreyer, which were used in QTL mapping by Tripathi et al. (2009a) and found to be in close linkage with a major QTL for of standard length. In addition, two newly designed primer pairs for genotyping of M9 and M987 in a microsatellite multiplex panel are shown (M9\_indel and M987\_indel).

a.

Marker	AT	Primer sequence	F tail	Reference
Pret32	58	F tgtaaaacgacggccagtGACCCCGACCCACAGAGAC R CACACGGCATAACGCACAT	Fam	Watanabe et al.2003
Pret69	53	F gccgctctagaactagtgGGTTGCTGCATGTGGTC R CTAGCGTTCCTGCATTCTG	Vic	""
Pret77	53	F tgtaaaacgacggccagtCTGCCACTGACCCTTTATTC R TTTGTTGCCAGCATTAC	Fam	""
Hull70-2	53	F gcaggaaacagctatgacTCAGGGAAGTGGTAATGATAATG R GGGTAACGACAAGTCTATCTTCAA	Pet	van Oosterhout et al. 2006
G82	58	F gcaggaaacagctatgacATCACAGGGCGTCATAGA R GATAAGGTTAGGGTTAGGAA	Pet	Shen et al. 2007
G102	53	F tgtaaaacgacggccagtGCTGGCTTCAATCAACTG R GATAGAGGTGGGGAGC	Fam	""
G289	58	F gccgctctagaactagtgATTGGGATTGATGAGGTG R GTGTTCCAGCAGGTCACT	Vic	""
Pr92	58	F gcaggaaacagctatgacACCTGTGCAGCAAGAC R TGGGCTGCTTTGTGAAGT	Pet	Becher et al. 2002
Pr39	53	F tagaaggcacagtcgaggGGTAAGGACTGATGAATAGCTTG R TTAGGGCCGTGCTTTTTG	Ned	""
FAM tail	50	TGTA AAAACGACGGCCAGT		Tysklind 2009
VIC tail	50	GCCGCTCTAGAACTAGTG		
PET tail	50	GCAGGAAACAGCTATGAC		
NED tail	50	TAGAAGGCACAGTCGAGG		

b.

Marker	AT	Primer sequence	F tail	Reference
M987_indel	58	F tgtaaaacgacggccagtAGCGGTCATGCACTAACAAG R ACATCTGCTCTGCCACAAA	Fam	This study
M9_indel	58	F gccgctctagaactagtgAATCGTTTTCCACAGGAGGTG R AGTCGCTAAATGCTGGTGGT	Vic	
M9	56	F CTCATTCTGTGCTTCAACCTG R GAAATGGCCTCCAGTATCTCC	---	Dreyer et al. 2007
M30	56	F TTGGTAGACCGAGACGTGAAG R TCAGCCTTGACATGAGTTACG	---	""
M61	56	F CAGTGGTCAGCGTGAAACTC R TGGAGAAGGTACAGGTCAACG	---	""
M155	56	F GACGAGTTTCTTCCGTTTCG R CAATGTTCTTAACCAAATGTCCG	---	""
M380	56	F CATGATTAAGTCTATTACGCTGCAC R AGATGTGAGTGGCAACATGG	---	Tripathi et al. 2009
M987	56	F AGCGGTCATGCACTAACAAG R CTGCCAACGAACATAAATGC	---	""
M1064	56	F GCTGGAGAACATCCTKGTG R GGAACATGTGGACGTAGAG	---	Hoffmann & Dreyer 2009, direct submission

- c. Primer pairs designed for sequencing of candidate genes in the present study, including redesigned primer pairs for marker M1078, initially obtained from C. Dreyer.
- d. SNP genotypes and flanking regions for polymorphisms genotyped in all male

**C.**

<b>Marker</b>	<b>AT</b>	<b>Primer sequence</b>
<b>M1078/ IGF-1 new primers</b>	53	F AAACCAGCTTTCACCTTTACACG R TGGAAAGTGAAGTGAAGGAA
<b>GH-1</b>	53	F TTTTGACCCAGGAGCATTGT R GGAGGTTTGAGGCGTCAGTA
<b>GH-2</b>	56	F CTGCAGAAATACGCCACTCA R ACAATGCTCCTGGGTCAAAA
<b>GH-3</b>	failed	F TTCTTTTGGCAGATTATTTCA R TTAGCACTGAGAAGGCAGAG
<b>GH-4</b>	53	F GGAAGTGAAGACGGGAATCA R CATGCGTCATGAGGAACAAG
<b>TBC-1</b>	53	F TGC GGCTGTATTGATAAAAT R CAACGAAAAGT GAGCTAGGA
<b>TBC-2</b>	54	F CGGCTAAAGACAGAGTGGACA R TACAGTCCTTTGCACCTTGC
<b>IGF2</b>	54	F AGATCCGGACACCACTCAC R AACTTTTTGGCCCTCAGGAT
<b>HSP-1</b>	failed	F CGTCTTCGAGGACTTCTCTG R GACGTCACCTTCTCGATGTT
<b>HSP-2</b>	53	F TGGCCAAGAAGTAGGAAGAA R TCGGTTGCTTTCTAATCCAG
<b>Myostatin-1</b>	56	F TAAAAGCCTGGACAGACCAA R ACTCACATCTCGGGTTCCT
<b>Myostatin-2</b>	failed	F AGTTTGCACTACGGACACC R CAAGCTACAGACTTTATTTGTTGA
<b>Pit54-1</b>	56	F TGCAGTGAAGGGAGTCTAC R CTACAGTTCCATTGCCCTGA
<b>Pit54-2</b>	56	F CAGCAAATGGCTTTAGAAGG R GTGCGTAAGTGC GACTTCTT
<b>Prolactin</b>	53	F TCATCACAGTGGTGTACCTG R AGCGAAGGACTTTCAAGAAG
<b>JunB</b>	53	F CTCCACAAGATGAACCAGAT R GTGCGTTAGGACCTTCTGTT

d.

SNP ID	Single nucleotide polymorphism and flanking regions
M9_403	CTGCRAGKGAATAATTTGGCACCGTTTAGTTTTACGGTATTTTAAAGTTTTCCAGGCAC[G/A]TTGGTCCYGACATGTGATCACACCAGCAATTTAGCGACTTCTGAGCGACTCTACGGATCA
M1046_Dreyer	ACATCACCATCTCTACGCCCTGGGTACCAACAACATGTACCATTCCAGGGGG[G/T]/C]TGTGGTCAATYRCCGGGATCWGGACGTGCTGCACCGTCTCCGGCATCCTCTTCATCATCTA
M1046_2	GACTCCATGATNGCAGMTCTGTCTCGCCTCCATCTGCAGCTTGTCCCATGGCC[A/G]TTCAGCCGCTACATCACCATCTTCTAGCCTTGCCTGACCAACATCTGCACCATCCAGCG
M30_Dreyer	CTGGTGTAGGGGGGATCTTTATACCCTGCCAACGTCAAGAGTCCAAAGGGCAAGGTAG[G/T]AAGCAGAGAYTCTGAAGCAAAAACAATCTTCTCAGGTAATCTTGTAAATCTATT
GH1	ATGTTAATACTGTTAATACTCTGCATCTTTAAGATTGTCAGATGAAAAATTRGCATAAAAAATCTT[G/T]CATCAATCACTACTCTCCCTGTGAACCTGTACCTGAGAAAACGTACCCTGAGCTTGCAAGGCCAA
GH2_60	TTCCTCTGCCTTCTCAGTGTAACTGCAGAAAATACGCCACTCAITTCGAAATAACCAATCA[A/G]AACCAAGGAGGAGGSTCTTAAcGCTGTCAATCAAGCGHATGTAAACACTCCCTCAAACCTCCCCYtGTCTCTCAACT
GH2_74	CTGCCTTCTCAGTCTAACTGCAGAAAATACGCCACTTCTGAAATAACCAATCARAACCAAGGAGG[G/C]/G]TCTTAAcGCTGTCAATCAAGCGHATGTAAACACTCCCTCAAACCTCCCCYtGTCTCTCAACT
GH2_165	TAAACACTCCTCAACCCCTCCCTCCCTCAACTAVGNACAGCTAGCACAGCCCTGTCAAGAA[G/T]GCTAATGTCTAGTATGATGACGRCAGATAAATAGTTTTCTGTAAITGTAAGT
GH2_211	AGCTAGCACAGCCTGTCATGAAGCTAATGCTAGTTAGCATGACGRCAGATAAATAGTTTTCTGTAA[A/C]TGTAAAGTTTCTCTRCCAATTAGCAGTGNKAGATGAGCATAAATTTGCTAGCACTAAGACCC
Myostatin	CgITTTGGAGACCCGTTGGAGTCAAGTTTCAACCCAGCTGAGCCGCCAAMAWTGAAGCTCCAGAA[A/C]AGTAACAGACCCGCTGGTTAGCAGCATYCCCCCKTTAGTcGACTAATAATTCAGCTAAACTCTC
Prolactin_1	GTGCCATGTGCAGAGCCGTGCCCTTCTCCTAGGACTCTACTTCTCCTTACGGTAGGATGCA[A/G]CGCTCCGACAAAATGCACCTCTGAGCAGCTGCCACCCAGGAGATGGTAAGACCTTMTGCTACCTGATT
Prolactin_2	CATTGTGATGATCACCCTTTTCTCCTAGGACTCTACTTCTCCTTACGGTAGGATGCA[A/G]ATGCCTCTCTGCCGACTGCCACACCTCCGCTCTACAGAGCCCAATGACAAAAGACAAGCTC
TBC1	AACTACTTATTTACCTGGAAGGGCTTCGGGAATATCTTACGGATTTTGGACTAGAGTAAGT[G/T]AATGTTAAGGGAGRCAGCGAGAGCAGCTCCCGTCCCGACTCCCAACATGGA
SF1	AGACAAAGTGTCCGGATACCACTACGGTCTCACCTGCGAAAAGCTCAAGGTGGTCAAGCTCTATAGCTC[A/T]CATTTNAGCCCGCACCGTCCGCATAGWTTCTGGGTCTAAATGGTCTTCAGATT



Hull170-2	F0	F1	F2	F3_S1	F3_S2	F3_C	F3_L1	F3_L2	F4_S1	F4_S2	F4_C	F4_L1	F4_L2	F5_S1	F5_S2	F5_C	F5_L1	F5_L2	F6_S1	F6_S2	F6_C	F6_L1	F6_L2
210	5.43	1.56	6.54	1.92	5.56	3.19	12.24	11.00	4.00	9.00	6.25	10.20	20.41		13.95	9.30	15.22	19.39		15.00	5.00	15.00	14.00
218	1.09		0.26																				
222	6.52	7.81	5.50	4.81	1.11	1.06	8.16	10.00	1.00	1.00		7.14	5.10	7.61		11.96	5.10	7.00	7.00	1.00		8.00	2.00
228	6.52	9.38	13.35	14.42	11.11	17.02	12.24	22.00	22.00	13.00	16.67	16.33	21.43	21.74	12.79	11.63	9.78	16.33	29.00	23.00	14.00	9.00	22.00
232	22.83	17.19	20.94	25.96	30.00	25.53	20.41	23.00	30.00	25.00	29.17	19.39	19.39	27.17	22.09	34.88	20.65	29.59	22.00	16.00	25.00	21.00	26.00
236	6.52	7.81	10.47	12.50	14.44	10.64	13.27	4.00	6.00	10.00	16.67	16.33	2.04	8.70	6.98	11.63	13.04	4.08	5.00	4.00	19.00	11.00	13.00
240	10.87	7.81	7.07	0.96	8.89	9.57	5.10	3.00	4.00	4.00	2.08	6.12	2.04	2.17	2.33	4.65	5.43	1.02	1.00	6.00	1.00	7.00	3.00
244	2.17	6.25	3.14	3.85	2.22	7.45	1.02	1.00	1.00	7.00	7.29	3.06	4.08	1.09	1.16	5.81	1.09	3.06	3.00	1.00	8.00	11.00	3.00
248	7.61	6.25	2.36	3.85	2.22	2.13	5.10	1.00	4.00	5.00	1.04	1.02	1.02	3.26	3.49	1.16	1.09	3.06	3.00	1.00	1.00	11.00	3.00
252	3.26	6.25	3.40		3.33	5.32	5.10	2.00	4.00	1.00	3.13	1.02	1.02	2.17	1.16	2.33	1.09	3.06	3.00	3.00	1.00	3.00	3.00
256	1.09																						
260		3.13	0.52			1.06							1.02		1.16								
264	2.17	3.13	4.19	1.92	3.33	3.19	2.04	2.00	1.00	4.00	3.13	5.10	2.04	5.43	12.79	6.98		1.02	4.00	12.00	10.00	2.00	3.00
266	1.09		2.62	1.92	1.11			1.00	1.00	1.00	1.04		2.04	1.09	2.33	2.33		5.10	2.00		1.00		
270	2.17	1.56	1.57		2.22	2.13	1.02	1.00	1.00	5.00	1.04				2.33	3.49	1.09		1.00	5.00	7.00	1.00	
274	1.09	1.56	1.31	0.96		1.06	1.02	1.00	3.00	3.00	1.04			1.16	1.16	1.16			3.00	2.00			
278	8.70	7.81	7.07	10.58	12.22	5.32	4.08	4.00	9.00	6.00	2.08	4.08	4.08	8.70	9.30	1.16	9.78	1.02	9.00	7.00	4.00	4.00	6.00
282	3.26	1.56	3.93	7.69		1.06	5.10	5.00	5.00	5.00	2.08	4.08	9.18	3.26	2.33		1.09	8.16	4.00	3.00	3.00	4.00	2.00
286	4.35	3.13	2.36	0.96		2.13	4.08	6.00	2.00	3.00	8.33	3.06	2.04	4.35	3.49	2.33	5.43	1.02	3.00		3.00	4.00	2.00
290			0.26		1.11	1.06				1.00		3.06			1.16	1.16							
292	2.17	6.25	2.88	7.69	1.11	1.06	2.04	4.00	5.00	2.00		3.06		3.26			3.26		5.00		3.00	3.00	1.00
296	1.09	1.56	0.26				3.06	1.00	4.00				3.06		1.16		1.09	2.04	2.00		1.00	1.00	2.00

Pret32	F0	F1	F2	F3_S1	F3_S2	F3_C	F3_L1	F3_L2	F4_S1	F4_S2	F4_C	F4_L1	F4_L2	F5_S1	F5_S2	F5_C	F5_L1	F5_L2	F6_S1	F6_S2	F6_C	F6_L1	F6_L2
123	1.11		1.31	0.96					3.00														
127	1.11																						
135	26.67	21.88	17.80	24.04	16.67	15.96	17.35	16.00	20.00	20.00	25.00	10.42	17.35	27.17	17.44	25.00	20.65	8.16	26.00	17.00	16.33	17.00	8.00
137	2.22		2.36	3.85	1.11	3.19		1.00	1.00	3.00				2.17	3.49				1.00	1.00			
139	7.78	3.13	4.97	5.77	8.89	9.57	3.06		1.00	6.00	6.25	2.08	2.04	1.09	10.47	5.95	1.09	1.02	8.00		9.18	1.00	
141	3.33	4.69	3.40	3.85	5.56	4.26	6.12	5.00	1.00	3.00	3.13	6.25	2.04	1.09	4.65	1.19		1.02	1.00	4.08	2.00		
145	5.56	1.56	2.62	1.92	3.33	2.13	6.12	7.00	3.00	5.00		3.13	1.02	11.96	4.65	2.38	5.43	5.10	6.00	14.00	9.18	7.00	9.00
147	11.11	10.94	12.83	10.58	10.00	19.15	15.31	14.00	12.00	11.00	17.71	12.50	13.27	13.04	8.14	21.43	17.39	13.27	20.00	7.00	15.31	11.00	14.00
149	5.56	4.69	3.93	3.85	6.67	7.45	4.08	4.00	4.00	7.00	12.50	7.29	4.08	3.26	9.30	3.57	6.52	2.04	3.00	13.00	7.14	8.00	
151	2.22	1.56	3.14	5.77	2.22	2.13	4.08	7.00	2.00		1.04	4.17	4.08	2.17	1.16		1.09	2.04	1.00	1.00	2.04	2.00	2.00
153	1.11																						
155	14.44	23.44	19.63	12.50	23.33	17.02	25.51	25.00	23.00	17.00	19.79	31.25	24.49	15.22	12.79	26.19	27.17	25.51	24.00	14.00	20.41	26.00	23.00
157	4.44	7.81	9.69	12.50	7.78	8.51	9.18	8.00	14.00	9.00	4.17	8.33	8.16	13.04	8.14	2.38	13.04	5.10	6.00	9.00	4.08	17.00	6.00
159	11.11	15.63	6.81	5.77	5.56	3.19	8.16	7.00	3.00	5.00	4.17	5.21	12.24	3.26	9.30	3.57	4.35	16.33	10.00		1.02	5.00	22.00
171	1.11																						
173		3.13	9.42	5.77	5.56	7.45	1.02	5.00	13.00	10.00	6.25	9.38	10.20	5.43	6.98	8.33	3.26	16.33	10.00	6.00	10.20	4.00	14.00
175	1.11		1.83	2.88	3.33			1.00	3.00					1.09	2.33				2.00		1.02		
177		1.56	0.26						1.00				1.02		1.16		4.08						2.00

G82	F0	F1	F2	F3_S1	F3_S2	F3_C	F3_L1	F3_L2	F4_S1	F4_S2	F4_C	F4_L1	F4_L2	F5_S1	F5_S2	F5_C	F5_L1	F5_L2	F6_S1	F6_S2	F6_C	F6_L1	F6_L2
184	48.91	54.69	49.74	47.12	47.78	56.38	65.31	49.00	51.00	45.00	68.75	56.12	50.00	58.70	39.53	70.93	59.78	51.02	57.00	40.00	64.00	64.00	47.00
186	43.48	29.69	42.67	46.15	43.33	35.11	30.61	43.00	38.00	51.00	18.75	39.80	45.92	36.96	56.98	18.60	38.04	42.86	40.00	57.00	20.00	35.00	49.00
188	7.61	10.94	7.07	6.73	7.78	8.51	4.08	8.00	10.00	2.00	12.50	4.08	4.08	4.35	3.49	10.47	2.17	6.12	3.00	1.00	16.00	1.00	4.00
190		4.69	0.52		1.11				1.00	2.00										2.00			

P192	F0	F1	F2	F3_S1	F3_S2	F3_C	F3_L1	F3_L2	F4_S1	F4_S2	F4_C	F4_L1	F4_L2	F5_S1	F5_S2	F5_C	F5_L1	F5_L2	F6_S1	F6_S2	F6_C	F6_L1	F6_L2
172	26.09	35.94	30.37	28.85	41.11	30.85	33.67	42.00	32.00	39.00	41.67	27.55	35.71	23.91	32.56	43.02	30.43	37.76	22.00	32.00	36.00	26.00	28.00
174	13.04	20.31	21.47	23.08	15.56	20.21	20.41	21.00	11.00	12.00	23.96	28.57	14.29	20.65	11.63	17.44	21.74	10.20	9.00	20.00	32.00	27.00	15.00
176	42.39	25.00	28.53	31.73	31.11	30.85	19.39	19.00	38.00	28.00	17.71	29.59	27.55	43.48	24.42	18.60	28.26	28.57	49.00	17.00	14.00	26.00	27.00
178	14.13	14.06	14.14	11.54	7.78	11.70	22.45	13.00	15.00	19.00	15.63	11.22	15.31	9.78	25.58	19.77	13.04	14.29	12.00	24.00	15.00	14.00	22.00
182	2.17	3.13	3.66	2.88	3.33	6.38	4.08	5.00	1.00	1.00	1.04	2.04	6.12	2.33	1.16		3.26	9.18		2.00	3.00	2.00	7.00
186	1.09	1.56	1.83	1.92	1.11				3.00	1.00		1.02	1.02	2.17	3.49		3.26		8.00	5.00		5.00	1.00
188	1.09																						

G289	F0	F1	F2	F3_S1	F3_S2	F3_C	F3_L1	F3_L2	F4_S1	F4_S2	F4_C	F4_L1	F4_L2	F5_S1	F5_S2	F5_C	F5_L1	F5_L2	F6_S1	F6_S2	F6_C	F6_L1	F6_L2
301	2.17		0.26	3.85	1.11	2.13	1.02	1.00	4.00		2.08	11.22	12.24	4.35				2.04	2.00				3.00
305	1.09		2.36	3.85	10.00	3.19	9.18	9.00	2.00	1.00	10.42	20.41	27.55	31.52	16.28	32.56	3.26	13.27	2.00	1.00	14.00	9.00	19.00
307	7.61	10.94	10.21	26.92	18.89	19.15	27.55	16.00	24.00	22.00	18.75	13.27	20.41	2.17	3.49	5.81	28.26	16.33	19.00	15.00	25.00	28.00	16.00
309	23.91	15.63	17.54	1.92	1.11	3.19	4.08	4.00	1.00		3.13	13.27	8.16	2.17	1.16	3.49	14.13	7.14	4.00	3.00	2.00	10.00	3.00
311	3.26	4.69	2.09	1.92	0.96	4.26	1.00	1.00		1.00	1.04		1.02		1.16	1.16				5.00	3.00		
313	6.52	4.69	1.57	0.96	2.22	4.26	1.00	1.00															
315	8.70	6.25	10.73	12.50	8.89	12.77	8.16	12.00	7.00	21.00	10.42	14.29	5.10	17.39	22.09	12.79	11.96	6.12	12.00	25.00	10.00	15.00	7.00
317	9.78	18.75	18.32	7.69	20.00	13.83	17.35	22.00	9.00	15.00	11.46	13.27	20.41	3.26	22.09	2.33	11.96	16.33	4.00	24.00	3.00	12.00	15.00
319	14.13	6.25	9.69	7.69	4.44	6.38	5.10	8.00	14.00	8.00	3.13	6.12	5.10	13.04	11.63	4.65	13.04	5.10	11.00	4.00	6.00	12.00	3.00
321	1.09	3.13	2.36	1.92	1.92	2.13	3.06	4.00	3.00		3.13	3.06	3.06	1.09		2.33	3.06	3.06	5.00		6.00	3.00	3.00
323	1.09	20.31	11.26	15.38	16.67	18.09	9.18	7.00	11.00	13.00	17.71	10.20	6.12	11.96	6.98	13.95	8.70	11.22	14.00	5.00	16.00	6.00	16.00
325	9.78	3.13	9.69	13.46	12.22	11.70	10.20	10.00	20.00	9.00	11.46	2.04	4.08	9.78	1.16	9.30	2.17	13.27	22.00	10.00	12.00	2.00	10.00
327	1.09	1.56	0.26					3.00															
329	9.78	3.13	3.66	3.85	3.33	3.19	4.08	3.00	5.00	10.00	7.29	9.18	5.10	5.43	13.95	11.63	6.52	6.12	5.00	8.00	9.00	6.00	5.00
347		1.56																					

**Expected heterozygosity per locus and population**

	<b>F0</b>	<b>F1</b>	<b>F2</b>	<b>F3_S1</b>	<b>F3_S2</b>	<b>F3_C</b>	<b>F3_L1</b>	<b>F3_L2</b>	<b>F4_S1</b>	<b>F4_S2</b>	<b>F4_C</b>	<b>F4_L1</b>	<b>F4_L2</b>
Pret69	0.8452	0.8507	0.8481	0.8469	0.8674	0.8677	0.8616	0.8616	0.8053	0.8525	0.8743	0.8445	0.8273
Pret77	0.6350	0.7103	0.6583	0.7114	0.6532	0.7163	0.5716	0.5248	0.7026	0.6701	0.6254	0.6829	0.6045
Hull70-2	0.9104	0.9320	0.9039	0.8749	0.8562	0.8794	0.9001	0.8723	0.8461	0.8921	0.8485	0.8900	0.8660
G102	0.7645	0.6458	0.6215	0.5659	0.7061	0.6360	0.5883	0.6499	0.6438	0.5984	0.5961	0.6514	0.5153
Pret32	0.8762	0.8611	0.8848	0.8893	0.8871	0.8861	0.8632	0.8731	0.8598	0.8935	0.8478	0.8539	0.8647
G82	0.5721	0.6081	0.5670	0.5659	0.5843	0.5575	0.4831	0.5743	0.5913	0.5420	0.4816	0.5304	0.5430
Pr92	0.7224	0.7579	0.7606	0.7556	0.7106	0.7591	0.7631	0.7313	0.7248	0.7261	0.7206	0.7494	0.7566
G289	0.8841	0.8854	0.8814	0.8613	0.8694	0.8810	0.8618	0.8859	0.8608	0.8519	0.8860	0.8761	0.8548

	<b>F5_S1</b>	<b>F5_S2</b>	<b>F5_C</b>	<b>F5_L1</b>	<b>F5_L2</b>	<b>F6_S1</b>	<b>F6_S2</b>	<b>F6_C</b>	<b>F6_L1</b>	<b>F6_L2</b>
Pret69	0.7907	0.8320	0.8695	0.8715	0.8452	0.8079	0.7711	0.8327	0.8471	0.8329
Pret77	0.7021	0.6862	0.6200	0.5936	0.6369	0.6774	0.5986	0.6014	0.6822	0.5673
Hull70-2	0.8581	0.8908	0.8386	0.8858	0.8405	0.8511	0.8778	0.8640	0.8931	0.8475
G102	0.6316	0.5902	0.5767	0.6992	0.5956	0.6602	0.5814	0.5539	0.7216	0.4786
Pret32	0.8576	0.9111	0.8184	0.8349	0.8582	0.8242	0.8907	0.8809	0.8547	0.8491
G82	0.5227	0.5239	0.4566	0.5029	0.5580	0.5194	0.5198	0.5301	0.4725	0.5428
Pr92	0.7088	0.7625	0.7190	0.7695	0.7442	0.6895	0.7760	0.7323	0.7772	0.7806
G289	0.8321	0.8462	0.8378	0.8505	0.8914	0.8691	0.8418	0.8626	0.8531	0.8780

**Observed heterozygosity per locus and population**

	<b>F0</b>	<b>F1</b>	<b>F2</b>	<b>F3_S1</b>	<b>F3_S2</b>	<b>F3_C</b>	<b>F3_L1</b>	<b>F3_L2</b>	<b>F4_S1</b>	<b>F4_S2</b>	<b>F4_C</b>	<b>F4_L1</b>	<b>F4_L2</b>
Pret69	0.8913	0.8438	0.8586	0.8846	0.8889	0.8478	0.8367	0.8200	0.8000	0.8200	0.8958	0.8980	0.8367
Pret77	0.5870	0.6250	0.6126	0.7115	0.6667	0.7447	0.6735	0.5800	0.6800	0.6600	0.6667	0.6735	0.5918
Hull70-2	0.9348	0.8438	0.8534	0.8654	0.8889	0.8298	0.8163	0.9200	0.8000	0.8400	0.8125	0.8571	0.8980
G102	0.6739	0.5938	0.6126	0.5769	0.6000	0.6383	0.5510	0.5200	0.7400	0.5600	0.5833	0.6939	0.5918
Pret32	0.8222	0.7188	0.8743	0.8077	0.8444	0.8085	0.7959	0.8600	0.7800	0.8000	0.8750	0.9167	0.8571
G82	0.6304	0.4375	0.4869	0.5962	0.6222	0.5319	0.4490	0.6000	0.5600	0.5400	0.3750	0.6735	0.6735
Pr92	0.8043	0.7188	0.7644	0.7692	0.6889	0.6596	0.6939	0.7000	0.6800	0.6800	0.7292	0.7347	0.8163
G289	0.7826	1.0000	0.8848	0.8462	0.8000	0.8085	0.8980	0.8600	0.8800	0.8200	0.8750	0.8163	0.7755

	<b>F5_S1</b>	<b>F5_S2</b>	<b>F5_C</b>	<b>F5_L1</b>	<b>F5_L2</b>	<b>F6_S1</b>	<b>F6_S2</b>	<b>F6_C</b>	<b>F6_L1</b>	<b>F6_L2</b>
Pret69	0.8043	0.8837	0.8140	0.8696	0.9388	0.8000	0.7600	0.9000	0.8600	0.8200
Pret77	0.7391	0.6977	0.6512	0.6087	0.6122	0.6600	0.6000	0.4600	0.7800	0.5600
Hull70-2	0.8696	0.8605	0.8140	0.9783	0.8776	0.8000	0.8400	0.9000	0.8000	0.8600
G102	0.6304	0.6977	0.5349	0.7174	0.6735	0.7800	0.6400	0.5800	0.6400	0.4600
Pret32	0.8043	0.8372	0.7143	0.7826	0.7959	0.8600	0.8600	0.8571	0.8400	0.8800
G82	0.5217	0.5814	0.4884	0.3261	0.4694	0.5600	0.4600	0.3800	0.5000	0.6400
Pr92	0.7391	0.8140	0.6744	0.8478	0.7347	0.7200	0.7800	0.8600	0.8600	0.7400
G289	0.8261	0.7907	0.9070	0.9348	0.9184	0.9000	0.9000	0.8200	0.7800	0.9800



## APPENDIX III – Full AMOVA Tables

- Summary of AMOVA statistics, showing the value of the different variance component among treatments ( $V_A$ ) among lines within treatments ( $V_B$ ) and within lines ( $V_C$ ) and the % of variation they explain for over all neutral microsatellite loci and in each of the generations F3-F6, as well as for a temporal hierarchy over these generations.
- Per-locus AMOVA tables showing sum of squares of the deviations (SSD), degrees of freedom (d.f.) the different variance components among treatments ( $V_A$ ) among lines within treatments ( $V_B$ ) and within lines ( $V_C$ ) and the % of variation they explain for each neutral microsatellite locus and in each of the generations F3-F6, as well as for a temporal hierarchy over these generations.
- Fixation indices among treatments ( $F_{CT}$ ) among lines within treatments ( $F_{SC}$ ) and within lines ( $F_{ST}$ ) and  $p$ -values for the analyses for each neutral microsatellite locus and over all loci for each of the generations F3-F6, as well as for a temporal analysis over these generations. Significant  $p$ -values are shown in bold.

a.

F3				F4			
Source of variation	Sum of squares	Variance component	Percentage of variation	Source of variation	Sum of squares	Variance component	Percentage of variation
Among treatments (Va, Fct)	10.614	0.01061	0.35	Among treatments (Va, Fct)	16.351	0.01747	0.58
Among lines within treatments (Vb, Fsc)	7.348	0.00665	0.22	Among lines within treatments (Vb, Fsc)	10.93	0.02504	0.83
Within lines (Vc, Fst)	1453.8	3.02417	99.43	Within lines (Vc, Fst)	1453.2	2.98568	98.60
Total	1471.7	3.04143		Total	1480.4	3.02819	

F5				F6			
Source of variation	Sum of squares	Variance component	Percentage of variation	Source of variation	Sum of squares	Variance component	Percentage of variation
Among treatments (Va, Fct)	19.107	0.00345	0.11	Among treatments (Va, Fct)	21.352	-0.01444	-0.48
Among lines within treatments (Vb, Fsc)	18.44	0.06804	2.24	Among lines within treatments (Vb, Fsc)	25.985	0.10043	3.31
Within lines (Vc, Fst)	1331.3	2.96701	97.65	Within lines (Vc, Fst)	1459	2.9492	97.17
Total	1368.9	3.0385		Total	1506.3	3.03519	

Temporal hierarchy F3-F6			
Source of variation	Sum of squares	Variance component	Percentage of variation
Among treatments (Va, Fct)	81.566	0.04271	1.41
Among lines within treatments (Vb, Fsc)	58.44	0.00948	0.31
Within lines (Vc, Fst)	5697.3	2.98153	98.28
Total	5837.3	3.03372	

**b.**

LOCUS	AMONG TREATMENTS				AMONG LINES WITHIN TREATMENT				WITHIN LINES				
	SSD	d.f.	Va	% variation	SSD	d.f.	Vb	% variation	SSD	d.f.	Vc	% variation	
F3	Pret69	1.6157	2	0.0024	0.5442	0.8913	2	0.0002	0.0363	206.1294	479	0.4303	99.4195
	Pret-77	1.8855	2	0.0050	1.5671	0.3241	2	-0.0016	-0.4956	152.7143	481	0.3175	98.9285
	Hull 70-2	1.9380	2	0.0025	0.5696	1.1611	2	0.0015	0.3287	210.8844	481	0.4384	99.1017
	G102	0.3571	2	-0.0022	-0.7127	1.0545	2	0.0022	0.6970	150.8415	481	0.3136	100.0156
	Pret-32	1.4047	2	0.0023	0.5255	0.6843	2	-0.0010	-0.2266	211.5653	481	0.4398	99.7011
	G82	0.9762	2	-0.0004	-0.1375	1.1016	2	0.0028	1.0073	132.9283	481	0.2764	99.1302
	Pr92	1.3943	2	0.0015	0.4001	0.9323	2	0.0010	0.2565	179.0438	481	0.3722	99.3435
	G289	1.0421	2	-0.0005	-0.1126	1.1985	2	0.0017	0.3824	209.6587	481	0.4359	99.7302

F4	Pret69	2.6915	2	0.0025	0.5762	1.9303	2	0.0055	1.2852	204.6729	487	0.4203	98.1386
	Pret-77	0.9726	2	-0.0007	-0.2093	1.1965	2	0.0027	0.8231	160.0971	487	0.3287	99.3862
	Hull 70-2	2.0060	2	0.0009	0.2145	1.7204	2	0.0043	0.9787	211.5054	487	0.4343	98.8067
	G102	0.2347	2	-0.0039	-1.3008	1.4747	2	0.0044	1.4653	146.3841	487	0.3006	99.8355
	Pret-32	1.7462	2	0.0031	0.7014	0.7917	2	-0.0004	-0.0846	209.5621	485	0.4321	99.3832
	G82	4.0746	2	0.0103	3.6513	0.8587	2	0.0016	0.5762	131.0362	487	0.2691	95.7725
	Pr92	1.9332	2	0.0024	0.6489	1.1801	2	0.0023	0.6030	179.0900	487	0.3677	98.7481
	G289	2.6922	2	0.0030	0.6693	1.7781	2	0.0046	1.0462	210.8123	487	0.4329	98.2846

F5	Pret69	3.0138	2	0.0002	0.0466	3.0102	2	0.0118	2.7262	188.9342	449	0.4208	97.2272
	Pret-77	1.6309	2	-0.0007	-0.2260	1.8767	2	0.0067	2.0285	145.3559	449	0.3237	98.1974
	Hull 70-2	2.1130	2	0.0002	0.0333	2.1021	2	0.0067	1.5395	193.6131	449	0.4312	98.4272
	G102	2.1261	2	0.0012	0.3804	1.8079	2	0.0065	2.0378	139.0351	449	0.3097	97.5817
	Pret-32	2.4200	2	0.0007	0.1687	2.2571	2	0.0076	1.7468	191.3185	447	0.4280	98.0845
	G82	3.8170	2	0.0065	2.3902	1.9825	2	0.0080	2.9428	115.3987	449	0.2570	94.6670
	Pr92	1.6773	2	-0.0020	-0.5191	2.2802	2	0.0084	2.2225	166.3178	449	0.3704	98.2966
	G289	2.3088	2	-0.0026	-0.6022	3.1231	2	0.0124	2.8341	191.3588	449	0.4262	97.7682

F6	Pret69	2.9210	2	-0.0016	-0.3893	3.4450	2	0.0131	3.1220	202.5400	495	0.4092	97.2673
	Pret-77	1.8920	2	0.0014	0.4439	1.4400	2	0.0041	1.2802	154.7800	495	0.3127	98.2759
	Hull 70-2	3.2130	2	0.0032	0.7178	2.1950	2	0.0066	1.4986	214.5100	495	0.4334	97.7836
	G102	1.2930	2	-0.0106	-3.4259	4.6850	2	0.0204	6.6028	148.2900	495	0.2996	96.8232
	Pret-32	1.8557	2	-0.0051	-1.1544	3.4850	2	0.0131	2.9968	211.9545	493	0.4299	98.1576
	G82	4.3460	2	0.0051	1.8737	2.7000	2	0.0109	3.9761	127.9400	495	0.2585	94.1502
	Pr92	2.8000	2	-0.0042	-1.0784	4.1400	2	0.0169	4.3636	185.9000	495	0.3756	96.7148
	G289	3.0310	2	-0.0027	-0.6096	3.8950	2	0.0152	3.4250	213.0800	495	0.4305	97.1846

Locus	AMONG LINES				AMONG LINES WITHIN GENERATIONS				WITHIN LINES				
	SSD	d.f.	Va	% variation	SSD	d.f.	Vb	% variation	SSD	d.f.	Vc	% variation	
Temporal hierarchy F3-F6	Pret69	11.5850	4	0.0059	1.3843	9.1422	15	0.0020	0.4591	802.2764	1910	0.4200	98.1566
	Pret-77	7.1866	4	0.0038	1.1601	5.1291	15	0.0002	0.0682	612.9473	1912	0.3206	98.7718
	Hull 70-2	10.2423	4	0.0052	1.1885	8.0538	15	0.0011	0.2411	830.5129	1912	0.4344	98.5705
	G102	6.9595	4	0.0034	1.0862	6.5498	15	0.0014	0.4369	584.5508	1912	0.3057	98.4769
	Pret-32	8.7678	4	0.0044	1.0011	7.5654	15	0.0008	0.1706	824.4005	1906	0.4325	98.8284
	G82	15.9286	4	0.0096	3.4883	4.1341	15	0.0001	0.0387	507.3033	1912	0.2653	96.4730
	Pr92	9.1944	4	0.0045	1.1793	8.6311	15	0.0021	0.5586	710.3516	1912	0.3715	98.2621
	G289	11.7023	4	0.0060	1.3607	9.2342	15	0.0019	0.4343	824.9098	1912	0.4314	98.2051

C.

		Fixation indices:					
		FSC	<i>p</i> -value	FST	<i>p</i> -value	FCT	<i>p</i> -value
F3	Pret69	0.00036	0.406	0.00580	0.064	0.00544	0.135
	Pret-77	-0.00504	0.832	0.01071	<b>0.022</b>	0.01567	0.137
	Hull 70-2	0.00331	0.165	0.00898	<b>0.007</b>	0.00570	0.134
	G102	0.00692	0.165	-0.00016	0.386	-0.00713	1.000
	Pret-32	-0.00228	0.766	0.00299	0.301	0.00526	0.132
	G82	0.01006	0.107	0.00870	0.079	-0.00137	0.461
	Pr92	0.00258	0.287	0.00657	0.121	0.00400	0.268
	G289	0.00382	0.164	0.00270	0.166	-0.00113	0.675
	All loci	0.00219	0.124	0.00567	<b>0.001</b>	0.00349	0.077

F4	Pret69	0.01293	<b>0.009</b>	0.01861	<b>0.000</b>	0.00576	0.273
	Pret-77	0.00821	0.074	0.00614	<b>0.060</b>	-0.00209	0.659
	Hull 70-2	0.00981	<b>0.019</b>	0.01193	<b>0.001</b>	0.00215	0.394
	G102	0.01446	<b>0.025</b>	0.00164	0.127	-0.01301	1.000
	Pret-32	-0.00085	0.568	0.00617	0.067	0.00701	0.064
	G82	0.00598	0.154	0.04227	<b>0.000</b>	0.03651	0.132
	Pr92	0.00607	0.144	0.01252	<b>0.013</b>	0.00649	0.329
	G289	0.01053	<b>0.022</b>	0.01715	<b>0.000</b>	0.00669	0.204
	All loci	0.00832	<b>0.000</b>	0.01404	<b>0.000</b>	0.00577	<b>0.047</b>

F5	Pret69	0.02727	<b>0.000</b>	0.02773	<b>0.000</b>	0.00047	0.596
	Pret-77	0.02024	<b>0.008</b>	0.01803	<b>0.001</b>	-0.00226	0.732
	Hull 70-2	0.01540	<b>0.002</b>	0.01573	<b>0.000</b>	0.00033	0.531
	G102	0.02046	<b>0.009</b>	0.02418	<b>0.001</b>	0.00380	0.533
	Pret-32	0.01750	<b>0.003</b>	0.01915	<b>0.000</b>	0.00169	0.461
	G82	0.03015	<b>0.031</b>	0.05333	<b>0.000</b>	0.02390	0.400
	Pr92	0.02211	<b>0.004</b>	0.01703	<b>0.001</b>	-0.00519	0.664
	G289	0.02817	<b>0.000</b>	0.02232	<b>0.000</b>	-0.00602	0.728
	All loci	0.02242	<b>0.000</b>	0.02353	<b>0.000</b>	0.00114	0.432

F6	Pret69	0.03110	<b>0.000</b>	0.02733	<b>0.000</b>	-0.00389	0.599
	Pret-77	0.01286	<b>0.024</b>	0.01724	<b>0.002</b>	0.00444	0.328
	Hull 70-2	0.01509	<b>0.002</b>	0.02216	<b>0.000</b>	0.00718	0.204
	G102	0.06384	<b>0.000</b>	0.03177	<b>0.000</b>	-0.03426	1.000
	Pret-32	0.02963	<b>0.000</b>	0.01842	<b>0.000</b>	-0.01154	0.930
	G82	0.04052	<b>0.003</b>	0.05850	<b>0.000</b>	0.01874	0.400
	Pr92	0.04317	<b>0.000</b>	0.03285	<b>0.000</b>	-0.01078	0.668
	G289	0.03404	<b>0.000</b>	0.02815	<b>0.000</b>	-0.00610	0.732
	All loci	0.03293	<b>0.000</b>	0.02833	<b>0.000</b>	-0.00476	0.777

Temporal hierarchy F3-F6	Pret69	0.00466	0.004	0.01843	<b>0.000</b>	0.01384	<b>0.000</b>
	Pret-77	0.00069	0.308	0.01228	<b>0.000</b>	0.01160	<b>0.000</b>
	Hull 70-2	0.00244	0.059	0.01430	<b>0.000</b>	0.01188	<b>0.000</b>
	G102	0.00442	<b>0.040</b>	0.01523	<b>0.000</b>	0.01086	<b>0.002</b>
	Pret-32	0.00172	0.186	0.01172	<b>0.000</b>	0.01001	<b>0.000</b>
	G82	0.00040	0.410	0.03527	<b>0.000</b>	0.03488	<b>0.000</b>
	Pr92	0.00565	<b>0.010</b>	0.01738	<b>0.000</b>	0.01179	<b>0.000</b>
	G289	0.00440	<b>0.004</b>	0.01795	<b>0.000</b>	0.01361	<b>0.000</b>
	All loci	0.00317	0.000	0.01720	0.000	0.01408	0.000

- d. Summary of AMOVA statistics, showing the value of the different variance component among treatments ( $V_A$ ) among lines within treatments ( $V_B$ ) and within lines ( $V_C$ ) and the % of variation they explain for over all putative candidate loci in generations F5-F6, as well as for an analysis over these two generations combined.
- e. Per-locus AMOVA tables showing sum of squares of the deviations (SSD), degrees of freedom (d.f.) the different variance components among treatments ( $V_A$ ) among lines within treatments ( $V_B$ ) and within lines ( $V_C$ ) and the % of variation they explain for each putative candidate locus in generations F5-F6, as well as for an analysis over these two generations combined
- f. Fixation indices among treatments among treatments ( $F_{CT}$ ) among lines within treatments ( $F_{SC}$ ) and within lines ( $F_{ST}$ ) and  $p$ -values for the analyses for each candidate locus and over all loci for generations F5-F6, as well as for an analysis over these two generations combined. Significant  $p$ -values are shown in bold.

d.

F5				F6			
Source of variation	Sum of squares	Variance component	Percentage of variation	Source of variation	Sum of squares	Variance component	Percentage of variation
Among treatments (Va, Fct)	38.215	0.09014	2.41	Among treatments (Va, Fct)	60.38	0.09014	2.41
Among lines within treatments (Vb, Fsc)	12.695	0.03002	0.80	Among lines within treatments (Vb, Fsc)	31.656	0.12393	3.31
Within lines (Vc, Fst)	1591.8	3.61627	96.78	Within lines (Vc, Fst)	1733.5	3.53319	97.29
Total	16.42.662	3.73643		Total	1825.5	3.74726	

F5 & F6			
Source of variation	Sum of squares	Variance component	Percentage of variation
Among treatments (Va, Fct)	89.278	0.12043	3.22
Among lines within treatments (Vb, Fsc)	58.685	0.05088	1.36
Within lines (Vc, Fst)	3325.2	3.57254	95.42
Total	3425.8	3.64225	

e.

LOCUS	AMONG TREATMENTS				AMONG LINES WITHIN TREATMENT				WITHIN LINES			
	SSD	d.f.	Va	% variation	SSD	d.f.	Vb	% variation	SSD	d.f.	Vc	% variation
M9_403	4.0692	2	0.0117	6.7578	0.7360	2	0.0023	1.3309	70.6233	443	0.1594	91.9114
M1046_Dreyer	0.2661	2	-0.0033	-1.4937	1.2010	2	0.0043	1.9456	95.7488	435	0.2201	99.5481
M1046_2	0.2862	2	0.0005	0.3624	0.1538	2	-0.0005	-0.4299	55.8368	443	0.1260	100.0675
M30_Dreyer	5.5171	2	0.0131	5.1327	1.8412	2	0.0076	2.9650	103.6372	441	0.2350	91.9023
GH1	0.9761	2	0.0030	1.2460	0.1458	2	-0.0018	-0.7658	101.9648	433	0.2355	99.5199
GH2_60	1.7270	2	0.0045	1.7817	0.4918	2	0.0000	-0.0021	105.6913	429	0.2464	98.2204
GH2_74	0.6436	2	0.0014	0.8658	0.2416	2	-0.0005	-0.2853	71.4639	439	0.1628	99.4195
GH2_165	1.3539	2	0.0007	0.2894	1.1587	2	0.0037	1.4894	108.1780	441	0.2453	98.2212
GH2_211	0.3252	2	-0.0011	-0.5271	0.6249	2	0.0013	0.6387	86.9254	437	0.1989	99.8884
Myostatin	0.5298	2	0.0007	0.3556	0.3265	2	-0.0004	-0.2070	88.0694	439	0.2006	99.8515
Prolactin_1	3.0380	2	0.0099	3.9988	0.2519	2	-0.0013	-0.5053	104.4578	439	0.2379	96.5065
Prolactin_2	2.0250	2	0.0068	2.6757	0.1111	2	-0.0021	-0.8470	108.8279	439	0.2479	98.1713
TBC_1	3.0419	2	0.0098	3.8522	0.2773	2	-0.0012	-0.4692	107.5705	439	0.2450	96.6170
SF1	0.9003	2	0.0012	0.9210	0.5697	2	0.0017	1.2990	56.8071	439	0.1294	97.7800
Pr39	6.0115	2	0.0137	3.6626	2.1050	2	0.0076	2.0431	157.9892	449	0.3519	94.2943
M987	6.4829	2	0.0144	4.8242	2.3783	2	0.0100	3.3379	123.0177	449	0.2740	91.8379
M9_Indel	1.0209	2	0.0033	3.1640	0.0808	2	-0.0007	-0.6324	44.9423	449	0.1001	97.4684

LOCUS	AMONG TREATMENTS				AMONG LINES WITHIN TREATMENT				WITHIN LINES			
	SSD	d.f.	Va	% variation	SSD	d.f.	Vb	% variation	SSD	d.f.	Vc	% variation
M9_403	5.8177	2	0.0147	7.1411	1.1530	2	0.0039	1.9157	91.7390	491	0.1868	90.9432
M1046_Dreyer	0.7763	2	0.0007	0.2962	0.5637	2	0.0006	0.2583	110.0774	491	0.2242	99.4455
M1046_2	0.3267	2	0.0004	0.2960	0.1893	2	-0.0005	-0.3535	71.9339	493	0.1459	100.0575
M30_Dreyer	11.2936	2	0.0238	9.3413	3.7486	2	0.0169	6.6268	104.5388	489	0.2138	84.0319
GH1	0.6824	2	-0.0054	-2.2999	2.4224	2	0.0099	4.1809	112.5781	487	0.2312	98.1190
GH2_60	3.3237	2	0.0053	2.1024	1.6419	2	0.0058	2.2949	118.9618	491	0.2423	95.6027
GH2_74	0.4949	2	0.0014	0.9480	0.0400	2	-0.0013	-0.8662	73.7696	491	0.1502	99.9181
GH2_165	2.8362	2	-0.0021	-0.8535	3.4980	2	0.0153	6.1435	115.7283	491	0.2357	94.7100
GH2_211	0.3999	2	-0.0067	-3.2427	2.4620	2	0.0107	5.1519	97.6484	481	0.2030	98.0908
Myostatin	0.2943	2	0.0002	0.0834	0.2447	2	-0.0007	-0.3689	93.8112	489	0.1918	100.2855
Prolactin_1	7.3361	2	0.0185	7.2693	1.4650	2	0.0050	1.9732	113.7692	493	0.2308	90.7575
Prolactin_2	1.7217	2	0.0035	1.4026	0.6400	2	0.0008	0.3111	118.1771	485	0.2437	98.2863
TBC_1	7.2086	2	0.0218	8.4554	0.2650	2	-0.0011	-0.4062	117.0084	493	0.2373	91.9507
SF1	0.3287	2	-0.0040	-3.1068	1.5777	2	0.0067	5.2817	61.0533	491	0.1243	97.8251
Pr39	10.1460	2	0.0163	4.3122	4.9400	2	0.0213	5.6467	168.1500	495	0.3397	90.0411
M987	7.0500	2	0.0032	1.0995	6.0200	2	0.0275	9.3871	129.7100	495	0.2620	89.5134
M9_Indel	0.3430	2	-0.0014	-1.9130	0.7850	2	0.0032	4.4615	34.8300	495	0.0704	97.4515

LOCUS	AMONG TREATMENTS				AMONG LINES WITHIN TREATMENT				WITHIN LINES			
	SSD	d.f.	Va	% variation	SSD	d.f.	Vb	% variation	SSD	d.f.	Vc	% variation
M9_403	9.4036	2	0.0141	7.3706	3.2171	7	0.0030	1.5863	162.3623	934	0.1738	91.0432
M1046_Dreyer	0.9874	2	0.0008	0.3430	1.8489	7	0.0005	0.2003	205.8262	926	0.2223	99.4567
M1046_2	0.5565	2	0.0006	0.4659	0.6031	7	-0.0005	-0.3896	127.7707	936	0.1365	99.9237
M30_Dreyer	15.7407	2	0.0227	8.8989	7.3801	7	0.0088	3.4599	208.1759	930	0.2239	87.6413
GH1	1.6224	2	0.0015	0.6343	2.6046	7	0.0015	0.6304	214.5430	920	0.2332	98.7353
GH2_60	4.5830	2	0.0062	2.4555	3.1563	7	0.0022	0.8805	224.6532	920	0.2442	96.6640
GH2_74	1.0241	2	0.0015	0.9370	0.5041	7	-0.0009	-0.5703	145.2335	930	0.1562	99.6333
GH2_165	3.9277	2	0.0041	1.6614	5.0121	7	0.0051	2.0266	223.9063	932	0.2402	96.3120
GH2_211	0.6961	2	-0.0004	-0.1813	3.1945	7	0.0028	1.3542	184.5738	918	0.2011	98.8271
Myostatin	0.2766	2	-0.0001	-0.0594	1.2116	7	-0.0002	-0.1249	181.8806	928	0.1960	100.1843
Prolactin_1	9.6514	2	0.0149	5.9344	2.5675	7	0.0014	0.5619	218.2270	932	0.2342	93.5037
Prolactin_2	3.2699	2	0.0043	1.7258	2.4446	7	0.0011	0.4406	227.0050	924	0.2457	97.8336
TBC_1	9.6015	2	0.0154	6.0242	1.2135	7	-0.0007	-0.2808	224.5788	932	0.2410	94.2566
SF1	0.8702	2	0.0003	0.1973	2.5152	7	0.0025	1.9110	117.8604	930	0.1267	97.8917
Pr39	15.0672	2	0.0207	5.5186	8.6197	7	0.0093	2.4723	326.1392	944	0.3455	92.0091
M987	10.9931	2	0.0129	4.3789	11.0518	7	0.0137	4.6677	252.7277	944	0.2677	90.9535
M9_Indel	1.0068	2	0.0009	1.0736	1.5406	7	0.0014	1.6357	79.7723	944	0.0845	97.2907

f.

		Fixation indices:					
		FSC	<i>p-value</i>	FST	<i>p-value</i>	FCT	<i>p-value</i>
F5	M9_403	0.0143	0.126	0.0809	<b>0.000</b>	0.0676	0.136
	M1046_Dreyer	0.0192	0.083	0.0045	0.174	-0.0149	0.794
	M1046_2	-0.0043	0.557	-0.0007	0.485	0.0036	0.393
	M30_Dreyer	0.0313	<b>0.015</b>	0.0810	<b>0.000</b>	0.0513	0.263
	GH1	-0.0078	0.702	0.0048	0.241	0.0125	0.202
	GH2_60	0.0000	0.290	0.0178	<b>0.024</b>	0.0178	0.199
	GH2_74	-0.0029	0.469	0.0058	0.201	0.0087	0.329
	GH2_165	0.0149	0.089	0.0178	<b>0.031</b>	0.0029	0.532
	GH2_211	0.0064	0.163	0.0011	0.246	-0.0053	0.671
	Myostatin	-0.0021	0.511	0.0015	0.437	0.0036	0.533
	Prolactin_1	-0.0053	0.561	0.0349	<b>0.006</b>	0.0400	0.135
	Prolactin_2	-0.0087	0.813	0.0183	0.094	0.0268	0.068
	TBC_1	-0.0049	0.558	0.0338	<b>0.005</b>	0.0385	0.203
	SF1	0.0131	0.091	0.0222	<b>0.023</b>	0.0092	0.402
	Pr39	0.0212	<b>0.003</b>	0.0571	<b>0.000</b>	0.0366	0.203
	M987	0.0351	<b>0.001</b>	0.0816	<b>0.000</b>	0.0482	0.137
	M9_Indel	-0.0065	0.578	0.0253	<b>0.013</b>	0.0316	0.066
All loci	0.0082	<b>0.001</b>	0.0322	<b>0.000</b>	0.0241	<b>0.001</b>	

F6	M9_403	0.0206	<b>0.041</b>	0.0906	<b>0.000</b>	0.0714	0.336
	M1046_Dreyer	0.0026	0.244	0.0055	0.162	0.0030	0.459
	M1046_2	-0.0036	0.506	-0.0006	0.455	0.0030	0.527
	M30_Dreyer	0.0731	<b>0.000</b>	0.1597	<b>0.000</b>	0.0934	0.329
	GH1	0.0409	<b>0.005</b>	0.0188	<b>0.006</b>	-0.0230	0.734
	GH2_60	0.0234	<b>0.035</b>	0.0440	<b>0.001</b>	0.0210	0.349
	GH2_74	-0.0087	0.885	0.0008	0.443	0.0095	0.135
	GH2_165	0.0609	<b>0.000</b>	0.0529	<b>0.000</b>	-0.0085	0.399
	GH2_211	0.0499	<b>0.002</b>	0.0191	<b>0.006</b>	-0.0324	1.000
	Myostatin	-0.0037	0.591	-0.0029	0.662	0.0008	0.604
	Prolactin_1	0.0213	0.061	0.0924	<b>0.000</b>	0.0727	0.202
	Prolactin_2	0.0032	0.329	0.0171	0.081	0.0140	0.267
	TBC_1	-0.0044	0.582	0.0805	<b>0.000</b>	0.0846	0.065
	SF1	0.0512	<b>0.002</b>	0.0218	<b>0.003</b>	-0.0311	0.801
	Pr39	0.0590	<b>0.000</b>	0.0996	<b>0.000</b>	0.0431	0.201
	M987	0.0949	<b>0.000</b>	0.1049	<b>0.000</b>	0.0110	0.463
	M9_Indel	0.0438	<b>0.001</b>	0.0255	<b>0.002</b>	-0.0191	0.738
All loci	0.0339	<b>0.000</b>	0.0571	<b>0.000</b>	0.0241	<b>0.036</b>	

F5 & F6	M9_403	0.0171	<b>0.017</b>	0.0896	<b>0.000</b>	0.0737	<b>0.016</b>
	M1046_Dreyer	0.0020	0.279	0.0054	0.162	0.0034	0.232
	M1046_2	-0.0039	0.721	0.0008	0.470	0.0047	0.088
	M30_Dreyer	0.0380	<b>0.000</b>	0.1236	<b>0.000</b>	0.0890	<b>0.031</b>
	GH1	0.0063	0.090	0.0127	<b>0.018</b>	0.0063	0.204
	GH2_60	0.0090	<b>0.042</b>	0.0334	<b>0.000</b>	0.0246	0.053
	GH2_74	-0.0058	0.833	0.0037	0.312	0.0094	<b>0.027</b>
	GH2_165	0.0206	<b>0.002</b>	0.0369	<b>0.000</b>	0.0166	0.118
	GH2_211	0.0135	<b>0.015</b>	0.0117	<b>0.011</b>	-0.0018	0.497
	Myostatin	-0.0013	0.613	-0.0018	0.676	-0.0006	0.482
	Prolactin_1	0.0060	0.158	0.0650	<b>0.000</b>	0.0593	<b>0.005</b>
	Prolactin_2	0.0045	0.262	0.0217	<b>0.014</b>	0.0173	<b>0.044</b>
	TBC_1	-0.0030	0.631	0.0574	<b>0.000</b>	0.0602	<b>0.002</b>
	SF1	0.0192	<b>0.004</b>	0.0211	<b>0.001</b>	0.0020	0.367
	Pr39	0.0262	<b>0.000</b>	0.0799	<b>0.000</b>	0.0552	<b>0.004</b>
	M987	0.0488	<b>0.000</b>	0.0905	<b>0.000</b>	0.0438	0.092
	M9_Indel	0.0165	<b>0.004</b>	0.0271	<b>0.000</b>	0.0107	0.177
All loci	0.0140	<b>0.000</b>	0.0458	<b>0.000</b>	0.0322	<b>0.000</b>	

**APPENDIX IV – Pairwise  $F_{ST}$  and  $G''_{ST}$  estimates per locus**

	All Microsatellites		Pret69		Pret77		Hull70-2		G102		Pret-32		G82		Pr92		G289		
	Fst	G''st	Fst	G''st	Fst	G''st	Fst	G''st	Fst	G''st	Fst	G''st	Fst	G''st	Fst	G''st	Fst	G''st	
F3	S1-S2	0.0027	0.0111	0.0054	0.0373	-0.0050	-0.0156	0.0043	0.0325	0.0085	0.0228	0.0023	0.0207	-0.0089	-0.0209	0.0047	0.0177	0.0062	0.0459
	L1-L2	0.0015	0.0058	-0.0045	-0.0327	-0.0031	-0.0070	0.0022	0.0197	0.0034	0.0089	-0.0081	-0.0616	0.0300	0.0637	-0.0001	-0.0004	0.0012	0.0094
	S1-C	0.0016	0.0068	0.0165	0.1155	-0.0066	-0.0232	0.0051	0.0415	-0.0026	-0.0065	0.0022	0.0196	0.0087	0.0197	-0.0088	-0.0366	-0.0034	-0.0264
	S2-C	-0.0012	-0.0052	0.0063	0.0479	-0.0031	-0.0099	-0.0016	-0.0118	-0.0037	-0.0113	-0.0030	-0.0265	0.0018	0.0042	-0.0014	-0.0052	-0.0049	-0.0395
	C-L1	0.0035	0.0141	-0.0003	-0.0023	0.0182	0.0507	0.0051	0.0466	-0.0080	-0.0206	0.0026	0.0204	0.0005	0.0009	0.0055	-0.0229	0.0035	-0.0269
	C-L2	0.0062	0.0257	-0.0043	-0.0314	0.0293	0.0767	0.0110	0.0887	0.0004	0.0013	0.0044	0.0364	0.0000	0.0001	0.0068	0.0266	0.0043	0.0369
F4	S1-L1	0.0097	0.0387	0.0077	0.0527	0.0117	0.0328	0.0078	0.0695	-0.0053	-0.0125	0.0081	0.0662	0.0432	0.0911	0.0098	0.0409	0.0025	0.0184
	S1-L2	0.0121	0.0491	0.0188	0.1292	0.0247	0.0649	0.0116	0.0915	0.0069	0.0175	0.0083	0.0699	-0.0081	-0.0187	0.0134	0.0523	0.0164	0.1302
	S2-L1	0.0062	0.0249	-0.0026	-0.0193	0.0040	0.0101	0.0082	0.0678	0.0041	0.0115	-0.0041	-0.0329	0.0337	0.0720	0.0177	0.0676	-0.0012	-0.0089
	S2-L2	0.0033	0.0137	0.0005	0.0042	0.0116	0.0279	0.0224	0.1649	-0.0007	-0.0022	-0.0024	-0.0200	-0.0098	-0.0233	0.0035	0.0125	-0.0014	-0.0118
	S1-S2	0.0080	0.0326	0.0117	0.0685	0.0078	0.0248	0.0055	0.0422	0.0141	0.0372	-0.0028	-0.0229	0.0133	0.0308	0.0011	0.0039	0.0159	0.1110
	L1-L2	0.0085	0.0331	0.0148	0.0898	0.0078	0.0220	0.0133	0.1092	0.0163	0.0391	0.0004	0.0027	-0.0006	-0.0012	0.0105	0.0423	0.0043	0.0324
F5	S1-C	0.0154	0.0605	0.0099	0.0619	0.0044	0.0131	0.0101	0.0664	0.0018	0.0047	0.0090	0.0620	0.0502	0.1087	0.0354	0.1279	0.0107	0.0847
	S2-C	0.0149	0.0584	0.0008	0.0057	-0.0007	-0.0019	-0.0003	-0.0021	-0.0038	-0.0095	0.0004	0.0030	0.1351	0.2773	0.0080	0.0290	0.0071	0.0546
	C-L1	0.0129	0.0511	0.0043	0.0308	0.0161	0.0467	0.0038	0.0289	-0.0008	-0.0020	0.0176	0.1179	0.0536	0.1084	0.0155	0.0584	0.0061	0.0514
	C-L2	0.0198	0.0739	0.0189	0.1266	0.0052	0.0134	0.0288	0.2017	-0.0039	-0.0088	0.0072	0.0499	0.0935	0.1918	0.0072	0.0277	0.0126	0.0979
	S1-L1	0.0133	0.0549	0.0256	0.1462	0.0145	0.0473	0.0129	0.0983	-0.0087	-0.0246	0.0051	0.0356	-0.0033	-0.0074	0.0176	0.0668	0.0317	0.2409
	S1-L2	0.0137	0.0530	0.0312	0.1700	0.0076	0.0221	0.0192	0.1335	0.0182	0.0432	-0.0014	-0.0100	-0.0004	-0.0009	0.0011	0.0041	0.0275	0.1935
F6	S2-L1	0.0108	0.0445	0.0174	0.1150	-0.0006	-0.0018	0.0015	0.0142	0.0107	0.0286	0.0103	0.0819	0.0150	0.0324	0.0209	0.0796	0.0119	0.0879
	S2-L2	0.0116	0.0449	0.0400	0.2499	0.0023	0.0064	0.0170	0.1409	0.0036	0.0083	0.0005	0.0043	-0.0035	-0.0075	-0.0063	-0.0242	0.0251	0.1717
	S1-S2	0.0291	0.1146	0.0347	0.1834	0.0278	0.0909	0.0132	0.1051	0.0233	0.0599	0.0095	0.0828	0.0591	0.1240	0.0405	0.1528	0.0357	0.2224
	L1-L2	0.0159	0.0631	0.0222	0.1556	0.0118	0.0307	0.0178	0.1295	0.0187	0.0527	0.0231	0.1513	-0.0023	-0.0050	0.0054	0.0221	0.0213	0.1644
	S1-C	0.0229	0.0840	0.0158	0.0930	0.0280	0.0827	0.0149	0.0981	0.0178	0.0450	0.0138	0.0856	0.0408	0.0801	0.0617	0.2154	0.0008	0.0047
	S2-C	0.0326	0.1234	0.0103	0.0692	0.0001	0.0004	0.0084	0.0623	0.0087	0.0208	0.0196	0.1463	0.1970	0.3861	0.0036	0.0137	0.0417	0.2640
F5	C-L1	0.0181	0.0680	-0.0019	-0.0145	0.0168	0.0427	0.0218	0.1581	0.0543	0.1506	0.0009	0.0049	0.0442	0.0854	0.0111	0.0433	0.0141	0.0899
	C-L2	0.0200	0.0745	0.0143	0.0997	-0.0012	-0.0031	0.0122	0.0755	0.0055	0.0133	0.0252	0.1573	0.0797	0.1631	0.0073	0.0273	0.0255	0.1885
	S1-L1	0.0180	0.0704	0.0259	0.1533	0.0715	0.0228	0.0191	0.1486	0.0089	0.0266	0.0045	0.0292	-0.0122	-0.0251	0.0100	0.0382	0.0109	0.0684
	S1-L2	0.0306	0.1186	0.0754	0.4132	0.0159	0.0482	0.0245	0.1621	-0.0014	-0.0036	0.0390	0.2760	-0.0025	-0.0054	0.0320	0.1172	0.0355	0.2570
	S2-L1	0.0245	0.0994	0.0155	0.1045	0.0080	0.0220	0.0113	0.1010	0.0725	0.2043	0.0164	0.1298	0.0586	0.1208	0.0079	0.0338	0.0226	0.1489
	S2-L2	0.0228	0.0915	0.0570	0.3517	0.0011	0.0033	0.0139	0.1031	0.0278	0.0682	0.0218	0.1899	0.0195	0.0428	0.0048	0.0196	0.0285	0.2177
F6	S1-S2	0.0397	0.1497	0.0391	0.1857	-0.0007	-0.0020	0.0181	0.1336	0.0656	0.1725	0.0304	0.2130	0.0442	0.0919	0.0791	0.2959	0.0429	0.2957
	L1-L2	0.0262	0.1017	0.0237	0.1482	0.0265	0.0706	0.0111	0.0860	0.0601	0.1504	0.0295	0.1989	0.0384	0.0779	0.0065	0.0296	0.0259	0.1923
	S1-C	0.0334	0.1257	-0.0002	-0.0014	0.0149	0.0415	0.0333	0.2340	0.0417	0.1059	0.0051	0.0342	0.0456	0.0960	0.1183	0.4082	0.0144	0.1077
	S2-C	0.0399	0.1489	0.0537	0.2706	0.0165	0.0414	0.0257	0.1992	0.0069	0.0160	0.0071	0.0625	0.1629	0.3437	0.0091	0.0370	0.0502	0.3397
	C-L1	0.0246	0.0973	0.0370	0.2302	0.0423	0.1182	0.0156	0.1287	0.0431	0.1191	0.0097	0.0735	0.0325	0.0653	0.0109	0.0444	0.0157	0.1112
	C-L2	0.0235	0.0862	0.0150	0.0894	0.0085	0.0204	0.0099	0.0688	0.0003	0.0007	0.0291	0.2157	0.0976	0.2107	0.0286	0.1173	0.0084	0.0649
F6	S1-L1	0.0262	0.1046	0.0277	0.1604	0.0226	0.0705	0.0383	0.3011	0.0141	0.0456	0.0127	0.0792	-0.0014	-0.0029	0.0487	0.1824	0.0333	0.2402
	S1-L2	0.0275	0.1015	0.0222	0.1234	0.0063	0.0167	0.0177	0.1176	0.0450	0.1044	0.0418	0.2553	0.0084	0.0179	0.0396	0.1494	0.0320	0.2516
	S2-L1	0.0287	0.1136	0.0107	0.0558	0.0213	0.0590	0.0220	0.1936	0.0635	0.1822	0.0115	0.0904	0.0883	0.1753	0.0076	0.0341	0.0294	0.1927
	S2-L2	0.0233	0.0853	0.0520	0.2626	-0.0010	-0.0023	0.0117	0.0855	0.0348	0.0740	0.0289	0.2222	0.0022	0.0047	0.0017	0.0077	0.0417	0.2962

Fst	M9_indel	M987	M9_403	M1046_Dreyer	M1046_2	M30_Dreyer	GH1	GH2_60	GH2_74	GH2_165	GH2_211	Myostatin	Prolactin1	Prolactin2	TBC_1	SF1	Pr39	All SNPs	All Candidates
F5	0.194	0.682	0.334	0.456	0.276	0.503	0.518	0.551	0.337	0.520	0.442	0.337	0.459	0.438	0.501	0.270	0.790	0.425	0.448
	0.172	0.534	0.349	0.444	0.271	0.448	0.467	0.488	0.313	0.480	0.402	0.392	0.469	0.492	0.482	0.253	0.690	0.411	0.420
	-0.005	0.081	0.023	0.002	-0.005	0.071	-0.004	0.006	-0.005	0.029	0.018	-0.006	0.002	-0.014	0.002	0.042	0.031	0.011	0.018
	-0.005	0.003	-0.006	0.035	-0.005	-0.005	-0.008	0.000	0.002	0.003	-0.003	-0.001	-0.009	-0.006	-0.009	0.000	0.015	-0.001	0.001
F6	-0.004	0.101	-0.014	-0.009	0.014	-0.007	0.008	-0.007	0.006	0.000	-0.004	-0.012	0.021	-0.013	0.093	-0.005	0.013	0.006	0.013
	0.009	0.003	0.031	-0.007	-0.005	0.113	-0.006	0.012	0.031	0.078	-0.003	0.004	-0.004	-0.010	0.035	0.087	0.015	0.025	0.022
	-0.001	0.013	0.072	0.024	-0.005	0.006	0.014	0.016	-0.011	-0.012	0.006	-0.012	0.017	0.021	0.091	-0.009	0.055	0.017	0.021
	0.018	0.037	0.038	-0.010	0.022	0.022	0.034	-0.004	0.001	0.006	-0.007	-0.004	0.004	0.048	0.078	-0.007	0.014	0.017	0.018
F5	0.019	0.172	0.080	0.004	-0.009	0.023	-0.009	0.028	0.007	-0.003	-0.009	-0.013	0.097	0.008	-0.008	0.005	0.093	0.015	0.036
	0.047	0.227	0.044	-0.004	-0.009	0.046	-0.002	0.002	-0.009	-0.009	-0.008	0.004	0.072	0.030	-0.010	-0.011	0.062	0.011	0.036
	0.043	0.038	0.205	0.050	-0.013	0.194	-0.001	0.081	0.031	0.072	0.036	0.001	0.044	0.011	0.002	0.106	0.126	0.061	0.065
	0.078	0.058	0.158	-0.010	0.000	0.234	0.013	0.037	0.000	0.024	0.010	0.034	0.025	0.034	-0.004	0.056	0.073	0.047	0.051
F6	0.067	0.174	0.041	0.014	0.002	0.137	0.089	0.053	-0.009	0.123	0.086	-0.004	-0.013	0.005	-0.010	0.084	0.108	0.046	0.063
	0.028	0.002	-0.009	-0.008	-0.009	-0.007	-0.009	-0.008	-0.007	-0.006	0.008	-0.008	0.047	-0.002	0.001	0.024	0.014	0.001	0.003
	-0.007	-0.007	0.031	-0.007	0.024	0.065	0.058	0.026	-0.002	-0.010	0.004	-0.010	0.103	-0.009	0.074	0.020	0.063	0.029	0.029
	0.048	0.147	-0.010	0.001	-0.005	0.348	-0.006	-0.006	-0.007	0.128	0.032	0.013	0.085	0.020	0.058	0.013	0.041	0.060	0.065
F6	0.066	0.169	0.150	-0.004	0.000	-0.008	0.012	0.056	-0.009	-0.005	0.005	-0.008	0.014	0.004	0.152	0.032	0.093	0.032	0.050
	0.000	0.118	0.137	-0.005	-0.005	-0.005	0.015	0.070	-0.005	-0.011	-0.009	0.003	-0.003	-0.012	0.226	-0.009	0.047	0.036	0.043
	0.085	0.202	0.039	-0.009	-0.002	0.054	0.004	-0.004	0.007	-0.004	-0.010	-0.013	0.211	0.018	0.007	-0.007	0.118	0.026	0.052
	0.010	0.148	0.031	-0.003	0.002	0.037	0.002	0.000	0.018	-0.011	0.007	-0.009	0.056	-0.007	0.042	0.013	0.092	0.013	0.032
F6	-0.008	0.038	0.168	0.023	-0.009	0.329	0.032	0.091	-0.001	0.169	0.087	-0.002	0.188	0.076	0.013	0.102	0.210	0.102	0.107
	0.015	0.024	0.154	0.040	-0.010	0.299	0.036	0.108	0.008	0.128	0.025	-0.009	0.042	0.028	0.054	0.018	0.130	0.077	0.077



G"st	M9_indel	M987	M9_403	M1046_Dreyer	M1046_2	M30_Dreyer	GH1	GH2_60	GH2_74	GH2_165	GH2_211	Myostatin	Prolactin1	Prolactin2	TBC_1	SF1	Pr39	All SNPs	All Candidates
Average Ho	0.194	0.682	0.334	0.456	0.276	0.503	0.518	0.337	0.520	0.442	0.337	0.459	0.438	0.501	0.270	0.790	0.425	0.448	
Average He	0.172	0.534	0.349	0.444	0.271	0.448	0.467	0.488	0.313	0.480	0.402	0.392	0.469	0.482	0.253	0.690	0.411	0.420	
S1-S2	-0.007	0.156	0.040	0.004	-0.006	0.133	-0.007	0.011	-0.008	0.057	0.031	-0.009	0.003	0.004	0.052	0.029	0.020	0.031	
L1-L2	-0.005	0.008	-0.007	0.061	-0.006	-0.009	-0.015	0.000	0.002	0.005	-0.004	-0.002	-0.019	-0.017	0.000	0.001	-0.001	0.002	
S1-C	-0.005	0.185	-0.022	-0.016	0.019	-0.013	0.015	-0.014	0.009	0.001	-0.006	-0.021	0.037	0.176	-0.006	0.025	0.011	0.023	
S2-C	0.013	0.007	0.054	-0.013	-0.006	0.210	-0.012	0.023	0.046	0.151	-0.006	0.006	-0.008	0.067	0.109	0.039	0.044	0.039	
C-L1	-0.001	0.030	0.098	0.040	-0.006	0.012	0.027	0.033	-0.015	-0.023	0.010	-0.021	0.034	0.174	-0.013	0.039	0.029	0.036	
C-L2	0.021	0.090	0.053	-0.018	0.029	0.042	0.064	-0.007	0.001	0.012	-0.012	-0.007	0.009	0.149	-0.010	0.035	0.029	0.032	
S1-L1	0.024	0.326	0.110	0.007	-0.012	0.046	-0.017	0.055	0.010	-0.005	-0.013	-0.021	0.180	-0.016	0.007	0.076	0.026	0.062	
S1-L2	0.056	0.447	0.062	-0.008	-0.013	0.087	-0.003	0.004	-0.014	-0.017	-0.013	0.006	0.133	-0.020	-0.015	0.075	0.019	0.062	
S2-L1	0.055	0.094	0.298	0.086	-0.017	0.350	-0.003	0.156	0.047	0.140	0.059	0.002	0.085	0.004	0.139	0.122	0.103	0.113	
S2-L2	0.094	0.150	0.235	-0.019	0.000	0.414	0.024	0.071	0.000	0.048	0.017	0.056	0.049	-0.008	0.072	0.098	0.080	0.089	
S1-S2	0.079	0.370	0.074	0.027	0.004	0.248	0.162	0.104	-0.013	0.231	0.149	-0.006	-0.022	0.010	0.107	0.105	0.079	0.109	
L1-L2	0.031	0.006	-0.012	-0.014	-0.013	-0.012	-0.017	-0.015	-0.009	-0.010	0.013	-0.013	0.092	0.002	0.034	0.006	0.002	0.005	
S1-C	-0.009	-0.012	0.055	-0.013	0.034	0.119	0.104	0.051	-0.002	-0.019	0.007	-0.015	0.189	0.137	0.028	0.048	0.049	0.051	
S2-C	0.056	0.322	-0.019	0.002	-0.006	0.577	-0.011	-0.012	-0.010	0.241	0.059	0.020	0.158	0.107	0.015	0.109	0.102	0.112	
C-L1	0.075	0.335	0.236	-0.008	0.000	-0.014	0.022	0.109	-0.012	-0.009	0.008	-0.012	0.028	0.279	0.045	0.101	0.054	0.086	
C-L2	0.000	0.239	0.217	0.009	-0.006	-0.009	0.028	0.135	-0.007	-0.020	-0.015	0.004	-0.007	0.402	-0.012	0.081	0.060	0.073	
S1-L1	0.099	0.387	0.057	-0.015	-0.003	0.099	0.008	-0.008	0.010	-0.007	-0.015	-0.021	0.376	0.035	-0.011	0.109	0.044	0.088	
S1-L2	0.012	0.289	0.045	-0.005	0.004	0.069	0.004	0.001	0.025	-0.021	0.012	-0.015	0.102	-0.014	0.018	0.069	0.022	0.055	
S2-L1	-0.008	0.094	0.265	0.043	-0.013	0.551	0.061	0.174	-0.001	0.311	0.151	-0.004	0.341	0.145	0.133	0.204	0.173	0.182	
S2-L2	0.016	0.061	0.246	0.073	-0.013	0.509	0.069	0.205	0.011	0.239	0.047	-0.016	0.078	0.056	0.022	0.148	0.131	0.131	

## APPENDIX V – Haplotypes for five Growth Hormone SNPs

```
(1...1050)GAAAAAATAGTTTTCTCTGCCTTCTCAGTGCTAACTGCAGAAATACGCCACT
CATTCTGAAATAACCAATCA[A/G]AACCAGGAGGAGG[C/G]TCTTAAcGCTGTCAATCA
AGCGtATGTAAACACTCCTCAACCCTCCCCYtGTCTCTCAACTAYGTNACAGCTAGCACAG
CCTGTCATGAA[G/T]GCTAATGCTAGTTAGCATGACGRCAGATAAATAGTTTTCTGTAA
[A/C]TGTAAGTTGTTTCTTRCCATTAGCAGTGNKKAGATGAGCATAATTGCTAGCACTA
AGACCCGCCTCCTGGCCCTGATTGGTTGTTTTGACCCAGGAGCATTGTATTTGTGCAGA
TCGCTAAAATAGCTCAGTGAAGAGGTAGAGGAGCGCGAGTTTTTTCAGATTTTCTGCCTC
ATAGCATTCTGTCACAACAGAGATTTTCAAATAAATGTTAATACTGTTAATACTCTGCA
TCTTTAAGATTGTCAGATGAAAATTRGCATAAAAATCTT[G/T]CATCAATCATCACTATCT
CCCTGTGAACCTGTTACCTGAGAACTGACCACCTGAGCTTGCAGGCCAATCAGGACGT
ACCAGAGATCTTTACTGACGCCTCAAACCTCCAGTTGGCACCTTACGGAAATTATTATCA
GAGTCTGGAAGCTGACGAGTCGCTGCGGAGAACCTACGAACTGCTGGCGTGCTTCAA
AAGGACATGAA
```

*Poecilia reticulata* Growth Hormone sequence showing the location of five SNPs in grey as they appear in the sequence: GH2\_60, GH2\_74, GH2\_165, GH2\_174, GH2\_211, GH1.

Haplotype frequencies of nine different haplotypes as they are present in generation F0, showing the five SNP genotypes as a single haplotype, its frequency and the minimal number of recombinations required to obtain the haplotype from one of the most frequent combination of two haplotypes (GGTCT and ACGAG). The location of recombination indicates between which two SNPs recombination must have occurred; 12 indicating between GH2\_60 and GH2\_74, 23 between GH2\_74 and GH2\_165, 34, between GH2\_165 and GH2\_211 and 45 between GH2\_211 and GH1. Most recombination events surround GH2\_74. Genotype frequencies within the F0 were not in Hardy-Weinberg equilibrium (Hardy-Weinberg probability exact test, 10.000 batches and 10.000 iterations per batch.  $F_{IS}$ : 0.0891,  $p$ -value: 0.000).

Haplotype	F0 frequency	Number of recombinations	Location of recombination
GGTCT	27.63	0	---
ACGAG	11.84	0	---
AGGAG	19.74	2	12-23
ACGCT	17.11	1	34
GGTCG	7.89	1	45
AGTCG	7.89	2	12-45
AGGCT	5.26	3	12-23-34
AGGAT	1.32	3	12-23-45
GGGCG	1.32	3	23-34-45

## **APPENDIX VI – Full tables for genotypic differentiation**

The table shows test *p*-values, significant Benjamini-Hochberg corrected *p*-values and standard errors for each locus and each population pair within generations F5 and F6 for tests of pairwise genotypic differentiation. An exact G-test was used in GenePop with the following Markov chain parameters: dememorisation number 1000, 10.000 batches and 10.000 iterations per batch. Within-treatment comparisons are shown in light grey, comparisons between selected lines and the Control in white and between selection lines in dark grey.

<b>Locus</b>	<b>Population pair</b>	<b>P-Value</b>	<b>B-H corr. sign. p-values</b>	<b>St. error</b>
TBC_1	F5_L2 & F5_L1	0.8748		0.00005
TBC_1	F5_S2 & F5_S1	0.3359		0.00014
TBC_1	F6_L2 & F6_L1	0.3722		0.00016
TBC_1	F6_S2 & F6_S1	0.8904		0.00005
TBC_1	F5_C & F5_S1	0.0014	0.0084	0.00001
TBC_1	F5_C & F5_S2	0.0551		0.00007
TBC_1	F5_C & F5_L1	0.0007	0.0048	0.00001
TBC_1	F5_C & F5_L2	0.0051	0.0245	0.00002
TBC_1	F6_C & F6_S1	0.0046	0.0225	0.00002
TBC_1	F6_C & F6_S2	0.0176		0.00005
TBC_1	F6_C & F6_L1	0.0001	0.0009	0.00000
TBC_1	F6_C & F6_L2	0.0000	0.0000	0.00000
TBC_1	F5_L1 & F5_S1	1.0000		0.00000
TBC_1	F5_L1 & F5_S2	0.3097		0.00013
TBC_1	F5_L2 & F5_S1	0.8801		0.00005
TBC_1	F5_L2 & F5_S2	0.4617		0.00014
TBC_1	F6_L1 & F6_S1	0.2413		0.00014
TBC_1	F6_L1 & F6_S2	0.1774		0.00014
TBC_1	F6_L2 & F6_S1	0.0258		0.00005
TBC_1	F6_L2 & F6_S2	0.0205		0.00005

<b>GH1</b>	F5_L2 & F5_L1	0.7487		0.00009
<b>GH1</b>	F5_S2 & F5_S1	0.4793		0.00013
<b>GH1</b>	F6_L2 & F6_L1	1.0000		0.00000
<b>GH1</b>	F6_S2 & F6_S1	0.0023	0.0126	0.00001
<b>GH1</b>	F5_C & F5_S1	0.2193		0.00012
<b>GH1</b>	F5_C & F5_S2	0.6037		0.00011
<b>GH1</b>	F5_C & F5_L1	0.1534		0.00012
<b>GH1</b>	F5_C & F5_L2	0.0443		0.00007
<b>GH1</b>	F6_C & F6_S1	0.0073	0.0321	0.00003
<b>GH1</b>	F6_C & F6_S2	0.5504		0.00013
<b>GH1</b>	F6_C & F6_L1	0.1560		0.00012
<b>GH1</b>	F6_C & F6_L2	0.1186		0.00011
<b>GH1</b>	F5_L1 & F5_S1	0.8686		0.00005
<b>GH1</b>	F5_L1 & F5_S2	0.4162		0.00014
<b>GH1</b>	F5_L2 & F5_S1	0.3889		0.00015
<b>GH1</b>	F5_L2 & F5_S2	0.1277		0.00011
<b>GH1</b>	F6_L1 & F6_S1	0.2679		0.00015
<b>GH1</b>	F6_L1 & F6_S2	0.0435		0.00007
<b>GH1</b>	F6_L2 & F6_S1	0.3439		0.00016
<b>GH1</b>	F6_L2 & F6_S2	0.0435		0.00007
<b>GH2_60</b>	F5_L2 & F5_L1	0.3344		0.00014
<b>GH2_60</b>	F5_S2 & F5_S1	0.2598		0.00013
<b>GH2_60</b>	F6_L2 & F6_L1	0.8770		0.00005
<b>GH2_60</b>	F6_S2 & F6_S1	0.0222		0.00005
<b>GH2_60</b>	F5_C & F5_S1	0.8537		0.00005
<b>GH2_60</b>	F5_C & F5_S2	0.1536		0.00011
<b>GH2_60</b>	F5_C & F5_L1	0.1455		0.00011
<b>GH2_60</b>	F5_C & F5_L2	0.4847		0.00012
<b>GH2_60</b>	F6_C & F6_S1	0.0736		0.00009
<b>GH2_60</b>	F6_C & F6_S2	0.5766		0.00014
<b>GH2_60</b>	F6_C & F6_L1	0.0098	0.0413	0.00003
<b>GH2_60</b>	F6_C & F6_L2	0.0046	0.0225	0.00002
<b>GH2_60</b>	F5_L1 & F5_S1	0.0668		0.00008
<b>GH2_60</b>	F5_L1 & F5_S2	0.0041	0.0206	0.00002
<b>GH2_60</b>	F5_L2 & F5_S1	0.2775		0.00012
<b>GH2_60</b>	F5_L2 & F5_S2	0.0337		0.00005
<b>GH2_60</b>	F6_L1 & F6_S1	0.5432		0.00013
<b>GH2_60</b>	F6_L1 & F6_S2	0.0028	0.0147	0.00002
<b>GH2_60</b>	F6_L2 & F6_S1	0.3691		0.00015
<b>GH2_60</b>	F6_L2 & F6_S2	0.0013	0.0083	0.00001

<b>GH2_74</b>	F5_L2 & F5_L1	0.3087		0.00019
<b>GH2_74</b>	F5_S2 & F5_S1	0.5862		0.00014
<b>GH2_74</b>	F6_L2 & F6_L1	0.8275		0.00006
<b>GH2_74</b>	F6_S2 & F6_S1	0.8621		0.00006
<b>GH2_74</b>	F5_C & F5_S1	0.2518		0.00016
<b>GH2_74</b>	F5_C & F5_S2	0.0808		0.00010
<b>GH2_74</b>	F5_C & F5_L1	1.0000		0.00000
<b>GH2_74</b>	F5_C & F5_L2	0.3240		0.00018
<b>GH2_74</b>	F6_C & F6_S1	0.3823		0.00017
<b>GH2_74</b>	F6_C & F6_S2	0.5979		0.00014
<b>GH2_74</b>	F6_C & F6_L1	0.8456		0.00007
<b>GH2_74</b>	F6_C & F6_L2	0.5539		0.00016
<b>GH2_74</b>	F5_L1 & F5_S1	0.2381		0.00017
<b>GH2_74</b>	F5_L1 & F5_S2	0.0480		0.00009
<b>GH2_74</b>	F5_L2 & F5_S1	0.8544		0.00006
<b>GH2_74</b>	F5_L2 & F5_S2	0.3509		0.00017
<b>GH2_74</b>	F6_L1 & F6_S1	0.2567		0.00018
<b>GH2_74</b>	F6_L1 & F6_S2	0.4463		0.00018
<b>GH2_74</b>	F6_L2 & F6_S1	0.1263		0.00014
<b>GH2_74</b>	F6_L2 & F6_S2	0.2478		0.00018
<b>GH2_165</b>	F5_L2 & F5_L1	0.2907		0.00015
<b>GH2_165</b>	F5_S2 & F5_S1	0.0588		0.00008
<b>GH2_165</b>	F6_L2 & F6_L1	0.6425		0.00012
<b>GH2_165</b>	F6_S2 & F6_S1	0.0002	0.0018	0.00000
<b>GH2_165</b>	F5_C & F5_S1	0.3459		0.00014
<b>GH2_165</b>	F5_C & F5_S2	0.0054	0.0253	0.00002
<b>GH2_165</b>	F5_C & F5_L1	1.0000		0.00000
<b>GH2_165</b>	F5_C & F5_L2	0.2400		0.00013
<b>GH2_165</b>	F6_C & F6_S1	1.0000		0.00000
<b>GH2_165</b>	F6_C & F6_S2	0.0002	0.0014	0.00000
<b>GH2_165</b>	F6_C & F6_L1	0.6254		0.00012
<b>GH2_165</b>	F6_C & F6_L2	1.0000		0.00000
<b>GH2_165</b>	F5_L1 & F5_S1	0.3920		0.00016
<b>GH2_165</b>	F5_L1 & F5_S2	0.0117	0.0484	0.00003
<b>GH2_165</b>	F5_L2 & F5_S1	0.8740		0.00005
<b>GH2_165</b>	F5_L2 & F5_S2	0.0625		0.00007
<b>GH2_165</b>	F6_L1 & F6_S1	0.5124		0.00014
<b>GH2_165</b>	F6_L1 & F6_S2	0.0000	0.0000	0.00000
<b>GH2_165</b>	F6_L2 & F6_S1	1.0000		0.00000
<b>GH2_165</b>	F6_L2 & F6_S2	0.0003	0.0022	0.00000

<b>GH2_211</b>	F5_L2 & F5_L1	0.4624		0.00017
<b>GH2_211</b>	F5_S2 & F5_S1	0.1068		0.00011
<b>GH2_211</b>	F6_L2 & F6_L1	0.2311		0.00015
<b>GH2_211</b>	F6_S2 & F6_S1	0.0024	0.0134	0.00001
<b>GH2_211</b>	F5_C & F5_S1	0.4803		0.00015
<b>GH2_211</b>	F5_C & F5_S2	0.4754		0.00014
<b>GH2_211</b>	F5_C & F5_L1	0.2210		0.00015
<b>GH2_211</b>	F5_C & F5_L2	0.7178		0.00011
<b>GH2_211</b>	F6_C & F6_S1	0.2943		0.00017
<b>GH2_211</b>	F6_C & F6_S2	0.0450		0.00007
<b>GH2_211</b>	F6_C & F6_L1	0.2741		0.00017
<b>GH2_211</b>	F6_C & F6_L2	1.0000		0.00000
<b>GH2_211</b>	F5_L1 & F5_S1	0.7187		0.00010
<b>GH2_211</b>	F5_L1 & F5_S2	0.0475		0.00007
<b>GH2_211</b>	F5_L2 & F5_S1	0.8547		0.00007
<b>GH2_211</b>	F5_L2 & F5_S2	0.1960		0.00014
<b>GH2_211</b>	F6_L1 & F6_S1	1.0000		0.00000
<b>GH2_211</b>	F6_L1 & F6_S2	0.0019	0.0109	0.00001
<b>GH2_211</b>	F6_L2 & F6_S1	0.2483		0.00016
<b>GH2_211</b>	F6_L2 & F6_S2	0.0761		0.00009
<b>M1046_Dreyer</b>	F5_L2 & F5_L1	0.0631		0.00009
<b>M1046_Dreyer</b>	F5_S2 & F5_S1	0.2883		0.00015
<b>M1046_Dreyer</b>	F6_L2 & F6_L1	0.7479		0.00010
<b>M1046_Dreyer</b>	F6_S2 & F6_S1	0.1086		0.00011
<b>M1046_Dreyer</b>	F5_C & F5_S1	0.6387		0.00012
<b>M1046_Dreyer</b>	F5_C & F5_S2	0.6216		0.00012
<b>M1046_Dreyer</b>	F5_C & F5_L1	0.1028		0.00010
<b>M1046_Dreyer</b>	F5_C & F5_L2	0.8720		0.00005
<b>M1046_Dreyer</b>	F6_C & F6_S1	0.6447		0.00012
<b>M1046_Dreyer</b>	F6_C & F6_S2	0.3716		0.00015
<b>M1046_Dreyer</b>	F6_C & F6_L1	0.4658		0.00015
<b>M1046_Dreyer</b>	F6_C & F6_L2	0.2768		0.00015
<b>M1046_Dreyer</b>	F5_L1 & F5_S1	0.2854		0.00015
<b>M1046_Dreyer</b>	F5_L1 & F5_S2	0.0283		0.00006
<b>M1046_Dreyer</b>	F5_L2 & F5_S2	0.7560		0.00009
<b>M1046_Dreyer</b>	F5_L2 & F5_S1	0.4553		0.00015
<b>M1046_Dreyer</b>	F6_L1 & F6_S1	0.8733		0.00005
<b>M1046_Dreyer</b>	F6_L1 & F6_S2	0.0696		0.00009
<b>M1046_Dreyer</b>	F6_L2 & F6_S1	0.4937		0.00016
<b>M1046_Dreyer</b>	F6_L2 & F6_S2	0.0238		0.00005

M1046_2	F5_L2 & F5_L1	0.4521		0.00016
M1046_2	F5_S2 & F5_S1	0.5140		0.00017
M1046_2	F6_L2 & F6_L1	0.8485		0.00007
M1046_2	F6_S2 & F6_S1	0.3462		0.00019
M1046_2	F5_C & F5_S1	0.1968		0.00014
M1046_2	F5_C & F5_S2	0.4739		0.00015
M1046_2	F5_C & F5_L1	0.4297		0.00013
M1046_2	F5_C & F5_L2	0.0831		0.00012
M1046_2	F6_C & F6_S1	0.0950		0.00011
M1046_2	F6_C & F6_S2	0.5672		0.00014
M1046_2	F6_C & F6_L1	0.3458		0.00016
M1046_2	F6_C & F6_L2	0.5673		0.00014
M1046_2	F5_L1 & F5_S1	0.7081		0.00010
M1046_2	F5_L1 & F5_S2	1.0000		0.00000
M1046_2	F5_L2 & F5_S1	1.0000		0.00000
M1046_2	F5_L2 & F5_S2	0.3697		0.00014
M1046_2	F6_L1 & F6_S1	0.4535		0.00018
M1046_2	F6_L1 & F6_S2	0.8485		0.00007
M1046_2	F6_L2 & F6_S1	0.3463		0.00019
M1046_2	F6_L2 & F6_S2	1.0000		0.00000
M30_Dreyer	F5_L2 & F5_L1	0.5861		0.00013
M30_Dreyer	F5_S2 & F5_S1	0.0098	0.0413	0.00003
M30_Dreyer	F6_L2 & F6_L1	0.7338		0.00010
M30_Dreyer	F6_S2 & F6_S1	0.0003	0.0022	0.00000
M30_Dreyer	F5_C & F5_S1	0.5425		0.00013
M30_Dreyer	F5_C & F5_S2	0.0002	0.0018	0.00000
M30_Dreyer	F5_C & F5_L1	0.1996		0.00012
M30_Dreyer	F5_C & F5_L2	0.0655		0.00008
M30_Dreyer	F6_C & F6_S1	0.0074	0.0321	0.00003
M30_Dreyer	F6_C & F6_S2	0.0000	0.0000	0.00000
M30_Dreyer	F6_C & F6_L1	0.8659		0.00006
M30_Dreyer	F6_C & F6_L2	0.6209		0.00013
M30_Dreyer	F5_L1 & F5_S1	0.0999		0.00010
M30_Dreyer	F5_L1 & F5_S2	0.0000	0.0000	0.00000
M30_Dreyer	F5_L2 & F5_S1	0.0258		0.00005
M30_Dreyer	F5_L2 & F5_S2	0.0000	0.0000	0.00000
M30_Dreyer	F6_L1 & F6_S1	0.0138		0.00004
M30_Dreyer	F6_L1 & F6_S2	0.0000	0.0000	0.00000
M30_Dreyer	F6_L2 & F6_S1	0.0339		0.00006
M30_Dreyer	F6_L2 & F6_S2	0.0000	0.0000	0.00000

M9_indel	F5_L1 & F5_S1	0.1351		0.00010
M9_indel	F5_L1 & F5_S2	0.0281		0.00005
M9_indel	F5_L2 & F5_S1	0.0201		0.00004
M9_indel	F5_L2 & F5_S2	0.0027	0.0145	0.00001
M9_indel	F6_L1 & F6_S1	0.0019	0.0109	0.00001
M9_indel	F6_L1 & F6_S2	1.0000		0.00000
M9_indel	F6_L2 & F6_S1	0.2269		0.00012
M9_indel	F6_L2 & F6_S2	0.1997		0.00009
M9_indel	F5_C & F5_S1	0.4825		0.00013
M9_indel	F5_C & F5_S2	0.2403		0.00012
M9_indel	F5_C & F5_L1	0.4220		0.00012
M9_indel	F5_C & F5_L2	0.1557		0.00009
M9_indel	F6_C & F6_S1	0.8200		0.00006
M9_indel	F6_C & F6_S2	0.0226		0.00004
M9_indel	F6_C & F6_L1	0.0076	0.0326	0.00002
M9_indel	F6_C & F6_L2	0.4540		0.00012
M9_indel	F5_L2 & F5_L1	0.5450		0.00009
M9_indel	F5_S2 & F5_S1	0.5116		0.00013
M9_indel	F6_L2 & F6_L1	0.0917		0.00006
M9_indel	F6_S2 & F6_S1	0.0064	0.0291	0.00002
M9_403	F5_L2 & F5_L1	0.5128		0.00014
M9_403	F5_S2 & F5_S1	0.1338		0.00012
M9_403	F6_L2 & F6_L1	1.0000		0.00000
M9_403	F6_S2 & F6_S1	0.0328		0.00006
M9_403	F5_C & F5_S1	0.8771		0.00005
M9_403	F5_C & F5_S2	0.0751		0.00009
M9_403	F5_C & F5_L1	0.0120	0.0492	0.00003
M9_403	F5_C & F5_L2	0.0431		0.00007
M9_403	F6_C & F6_S1	0.0458		0.00007
M9_403	F6_C & F6_S2	0.8869		0.00005
M9_403	F6_C & F6_L1	0.0001	0.0007	0.00000
M9_403	F6_C & F6_L2	0.0002	0.0015	0.00000
M9_403	F5_L1 & F5_S1	0.0156		0.00004
M9_403	F5_L1 & F5_S2	0.0000	0.0001	0.00000
M9_403	F5_L2 & F5_S1	0.0453		0.00007
M9_403	F5_L2 & F5_S2	0.0001	0.0010	0.00000
M9_403	F6_L1 & F6_S1	0.0230		0.00006
M9_403	F6_L1 & F6_S2	0.0000	0.0002	0.00000
M9_403	F6_L2 & F6_S1	0.0451		0.00008
M9_403	F6_L2 & F6_S2	0.0001	0.0007	0.00000

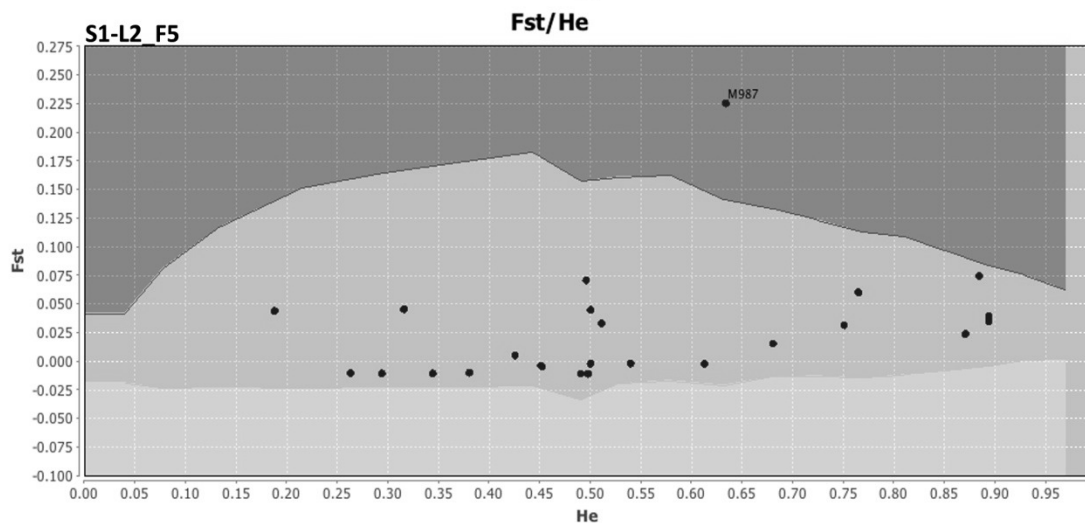
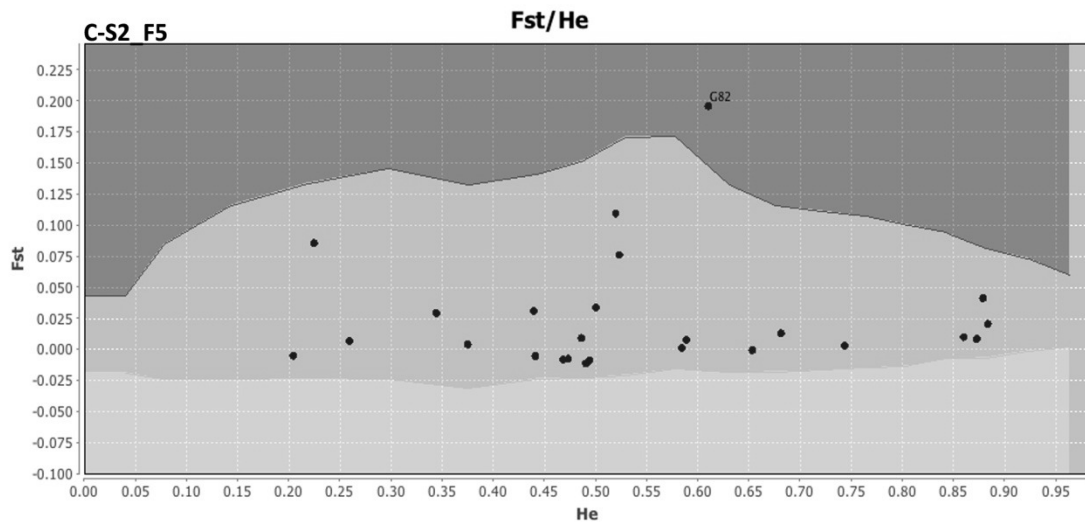


M987	F5_L2 & F5_L1	0.2068		0.00019
M987	F5_S2 & F5_S1	0.0002	0.0016	0.00000
M987	F6_L2 & F6_L1	0.1223		0.00012
M987	F6_S2 & F6_S1	0.0000	0.0000	0.00000
M987	F5_C & F5_S1	0.0001	0.0009	0.00000
M987	F5_C & F5_S2	0.1390		0.00016
M987	F5_C & F5_L1	0.0042	0.0210	0.00002
M987	F5_C & F5_L2	0.0014	0.0083	0.00001
M987	F6_C & F6_S1	0.9299		0.00006
M987	F6_C & F6_S2	0.0000	0.0000	0.00000
M987	F6_C & F6_L1	0.0000	0.0000	0.00000
M987	F6_C & F6_L2	0.0002	0.0015	0.00000
M987	F5_L1 & F5_S1	0.0000	0.0001	0.00000
M987	F5_L1 & F5_S2	0.0003	0.0022	0.00001
M987	F5_L2 & F5_S1	0.0000	0.0000	0.00000
M987	F5_L2 & F5_S2	0.0009	0.0058	0.00001
M987	F6_L1 & F6_S1	0.0000	0.0000	0.00000
M987	F6_L1 & F6_S2	0.0000	0.0000	0.00000
M987	F6_L2 & F6_S1	0.0000	0.0000	0.00000
M987	F6_L2 & F6_S2	0.0004	0.0030	0.00001
Myostatin	F5_L2 & F5_L1	0.3606		0.00016
Myostatin	F5_S2 & F5_S1	0.5186		0.00014
Myostatin	F6_L2 & F6_L1	0.6419		0.00012
Myostatin	F6_S2 & F6_S1	0.4836		0.00016
Myostatin	F5_C & F5_S1	0.7604		0.00009
Myostatin	F5_C & F5_S2	0.2995		0.00015
Myostatin	F5_C & F5_L1	1.0000		0.00000
Myostatin	F5_C & F5_L2	0.4401		0.00015
Myostatin	F6_C & F6_S1	0.6571		0.00012
Myostatin	F6_C & F6_S2	0.1710		0.00013
Myostatin	F6_C & F6_L1	0.6330		0.00012
Myostatin	F6_C & F6_L2	0.3392		0.00016
Myostatin	F5_L1 & F5_S1	0.8797		0.00005
Myostatin	F5_L1 & F5_S2	0.3091		0.00015
Myostatin	F5_L2 & F5_S1	0.3072		0.00016
Myostatin	F5_L2 & F5_S2	0.0530		0.00008
Myostatin	F6_L1 & F6_S1	1.0000		0.00000
Myostatin	F6_L1 & F6_S2	0.4552		0.00016
Myostatin	F6_L2 & F6_S1	0.6641		0.00012
Myostatin	F6_L2 & F6_S2	0.7651		0.00009

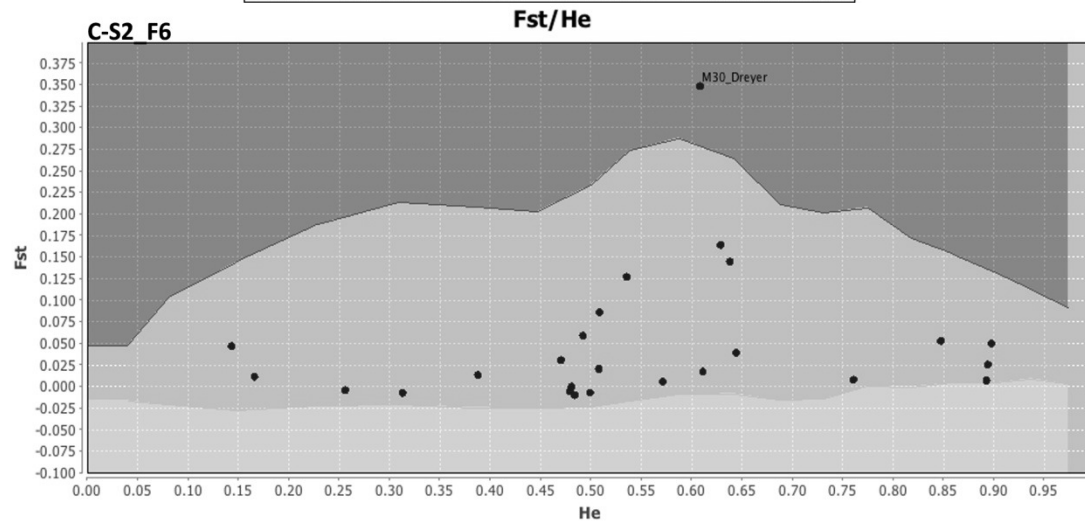
Pr39	F5_L2 & F5_L1	0.0262		0.00008
Pr39	F5_S2 & F5_S1	0.0014	0.0083	0.00001
Pr39	F6_L2 & F6_L1	0.0237		0.00008
Pr39	F6_S2 & F6_S1	0.0000	0.0000	0.00000
Pr39	F5_C & F5_S1	0.0575		0.00012
Pr39	F5_C & F5_S2	0.1685		0.00021
Pr39	F5_C & F5_L1	0.0003	0.0021	0.00001
Pr39	F5_C & F5_L2	0.2493		0.00023
Pr39	F6_C & F6_S1	0.0001	0.0010	0.00000
Pr39	F6_C & F6_S2	0.0000	0.0002	0.00000
Pr39	F6_C & F6_L1	0.0000	0.0000	0.00000
Pr39	F6_C & F6_L2	0.0005	0.0035	0.00001
Pr39	F5_L1 & F5_S1	0.0000	0.0000	0.00000
Pr39	F5_L1 & F5_S2	0.0000	0.0000	0.00000
Pr39	F5_L2 & F5_S1	0.0000	0.0002	0.00000
Pr39	F5_L2 & F5_S2	0.0002	0.0019	0.00001
Pr39	F6_L1 & F6_S1	0.0000	0.0000	0.00000
Pr39	F6_L1 & F6_S2	0.0000	0.0000	0.00000
Pr39	F6_L2 & F6_S1	0.0000	0.0000	0.00000
Pr39	F6_L2 & F6_S2	0.0000	0.0000	0.00000
SF1	F5_L2 & F5_L1	0.3387		0.00019
SF1	F5_S2 & F5_S1	0.0554		0.00007
SF1	F6_L2 & F6_L1	0.0948		0.00012
SF1	F6_S2 & F6_S1	0.0012	0.0076	0.00001
SF1	F5_C & F5_S1	0.4365		0.00015
SF1	F5_C & F5_S2	0.0034	0.0175	0.00002
SF1	F5_C & F5_L1	0.8426		0.00007
SF1	F5_C & F5_L2	0.6879		0.00012
SF1	F6_C & F6_S1	0.0850		0.00009
SF1	F6_C & F6_S2	0.1921		0.00011
SF1	F6_C & F6_L1	0.0523		0.00009
SF1	F6_C & F6_L2	1.0000		0.00000
SF1	F5_L1 & F5_S1	0.2544		0.00016
SF1	F5_L1 & F5_S2	0.0006	0.0042	0.00001
SF1	F5_L2 & F5_S1	0.8468		0.00006
SF1	F5_L2 & F5_S2	0.0156		0.00004
SF1	F6_L1 & F6_S1	0.7113		0.00013
SF1	F6_L1 & F6_S2	0.0009	0.0058	0.00001
SF1	F6_L2 & F6_S1	0.1532		0.00016
SF1	F6_L2 & F6_S2	0.1503		0.00013

## APPENDIX VII – Fdist figures

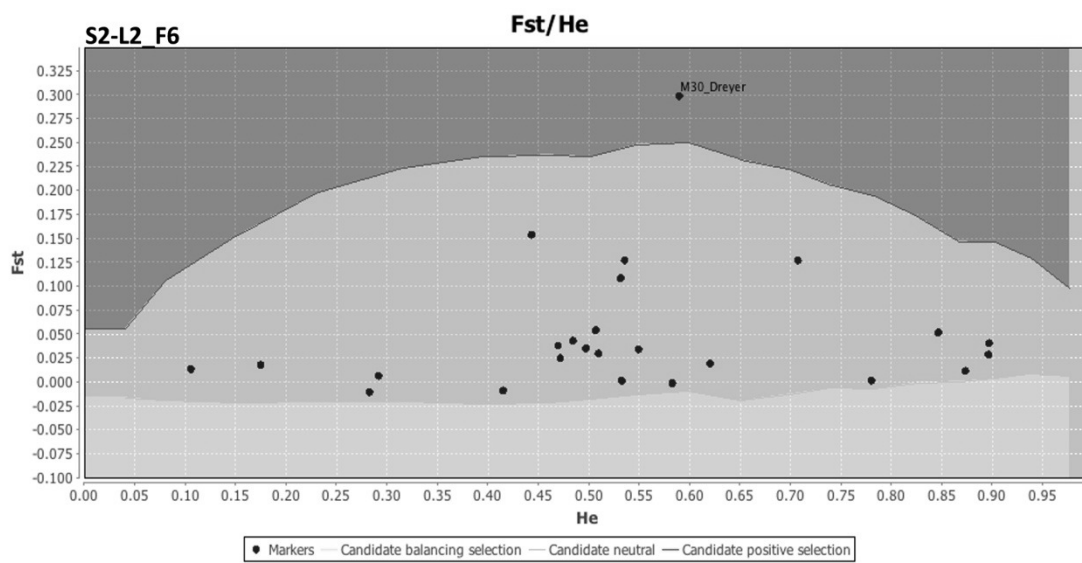
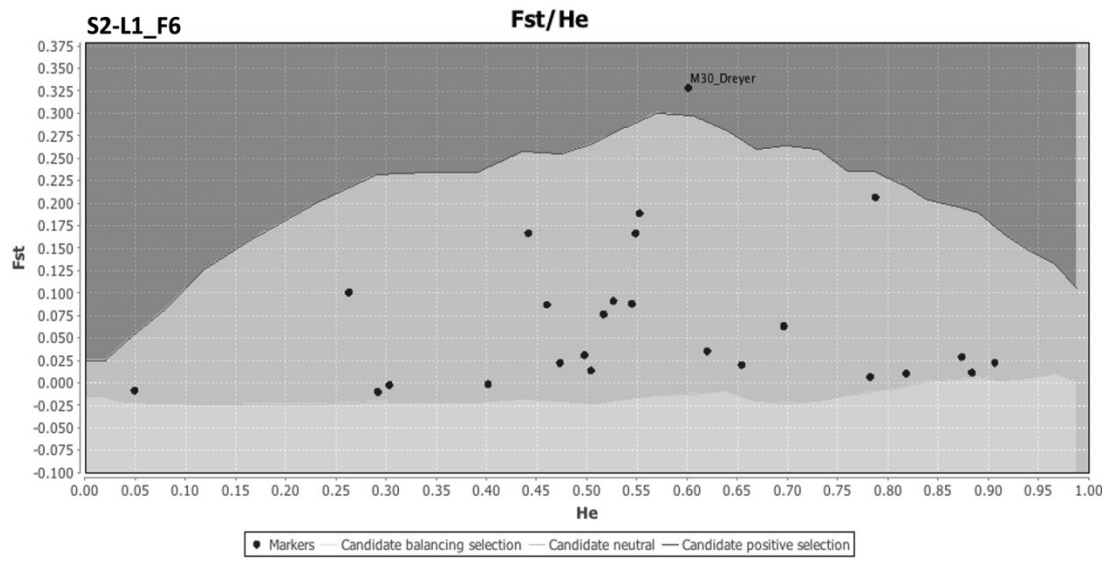
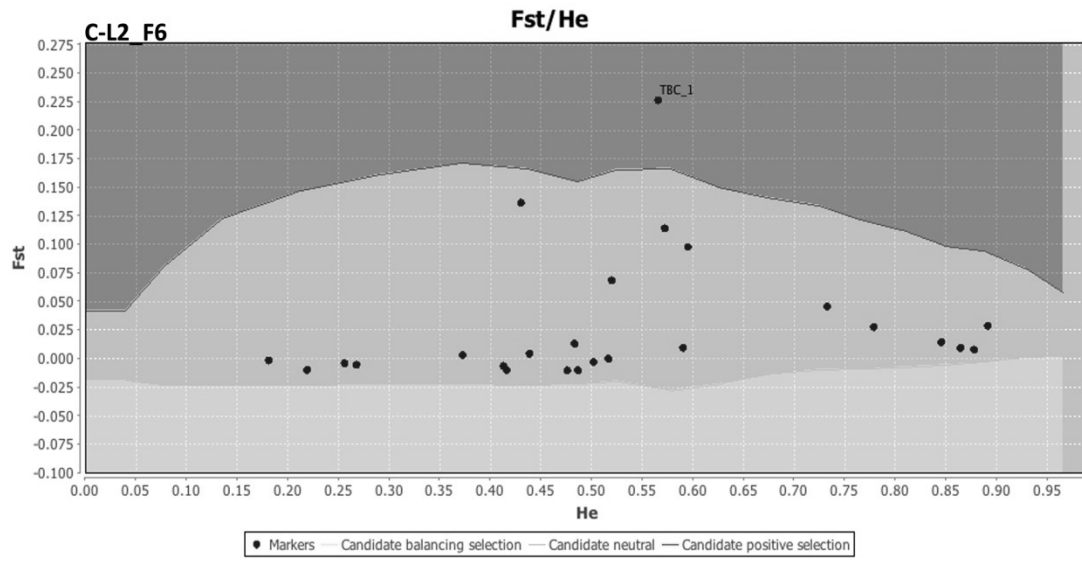
The figures show the results for pairwise comparisons in LOSITAN in which significant outlier loci were detected. Each graph plots  $F_{ST}$  as a function of marker heterozygosity. Candidates for divergent selection lie in the dark grey area above the simulated neutral distribution (light grey). Marker names are shown for outlier loci only.

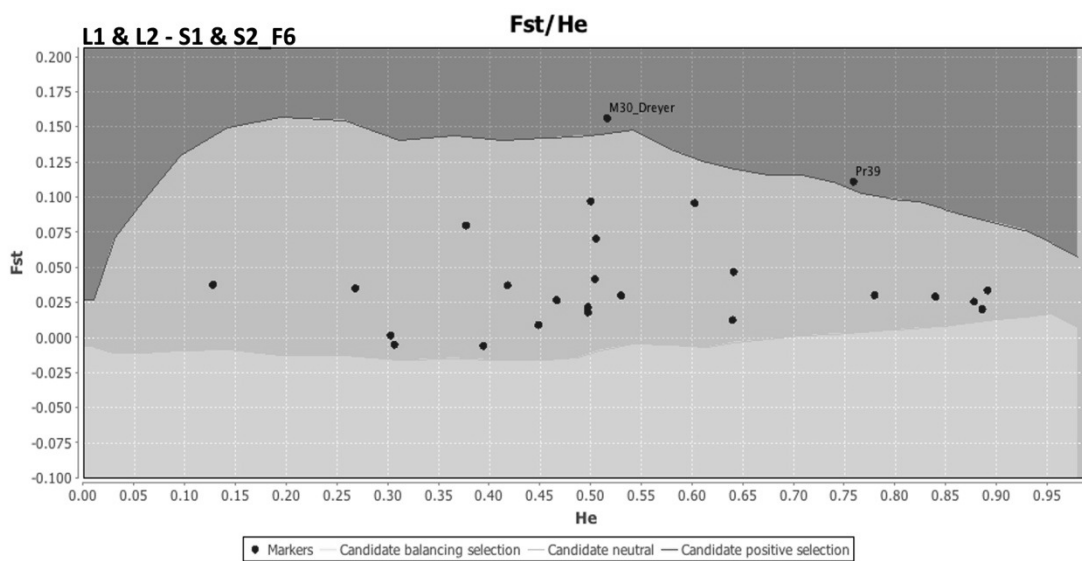
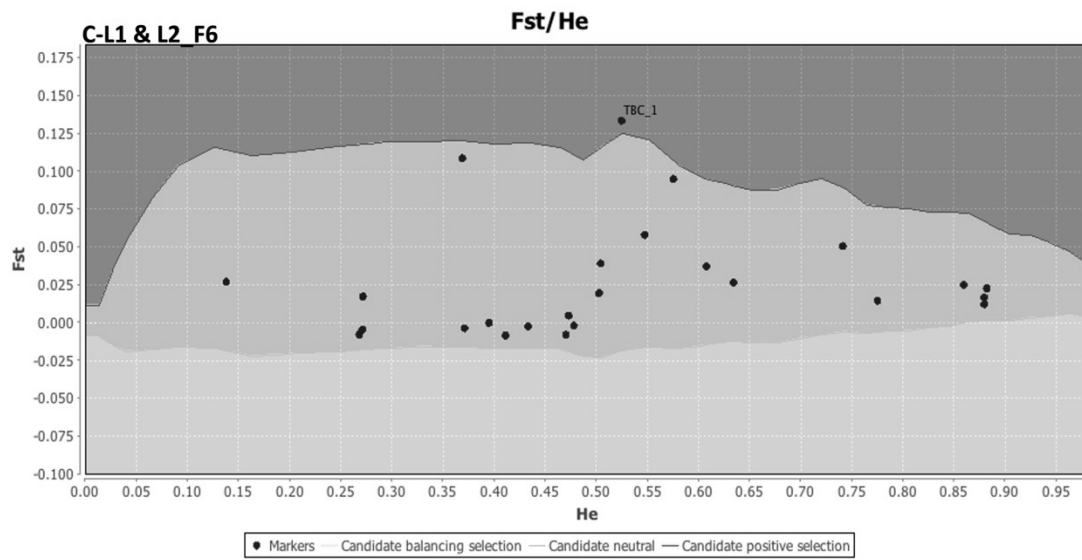
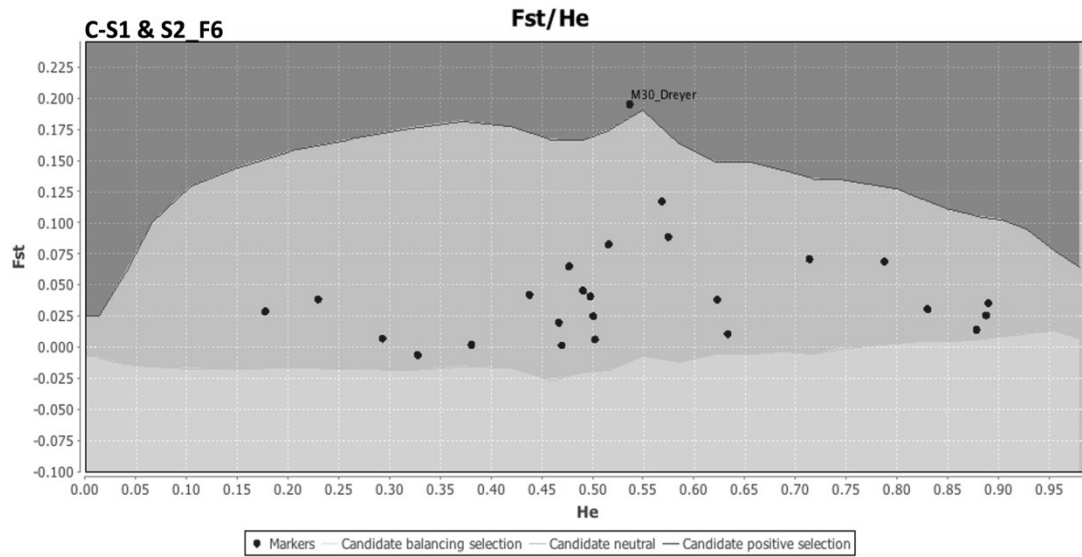


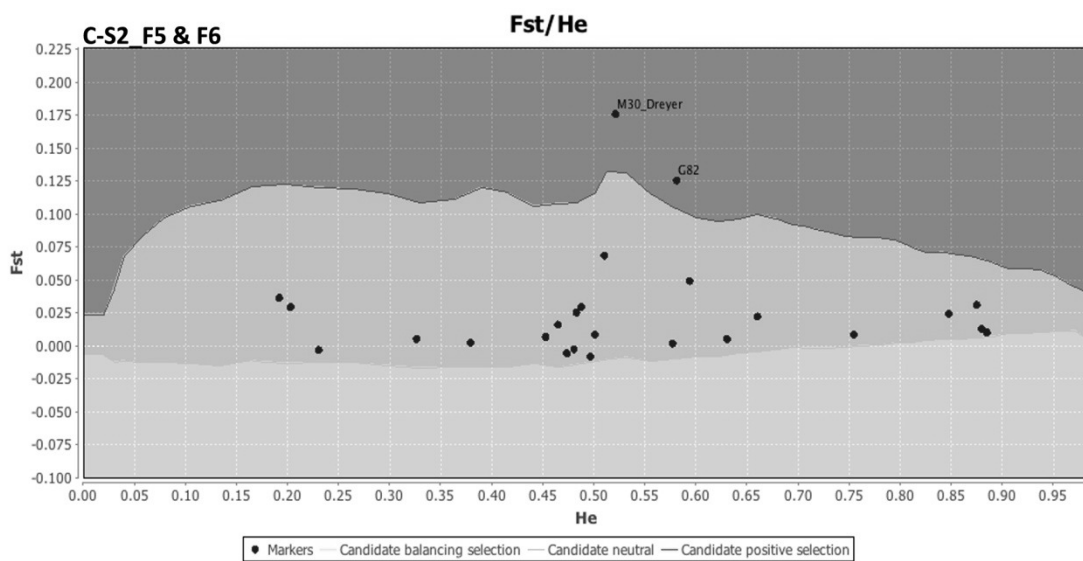
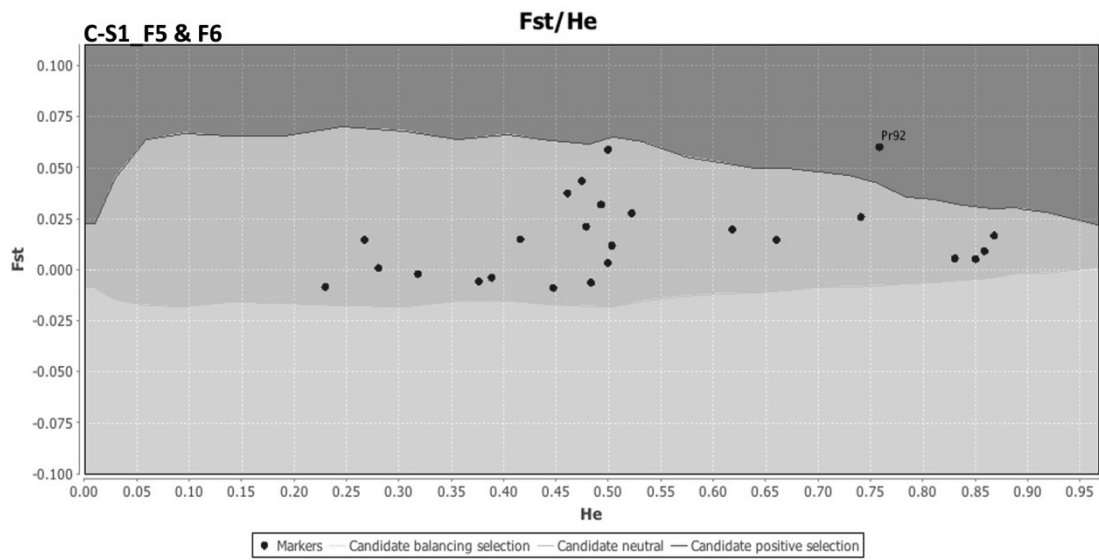
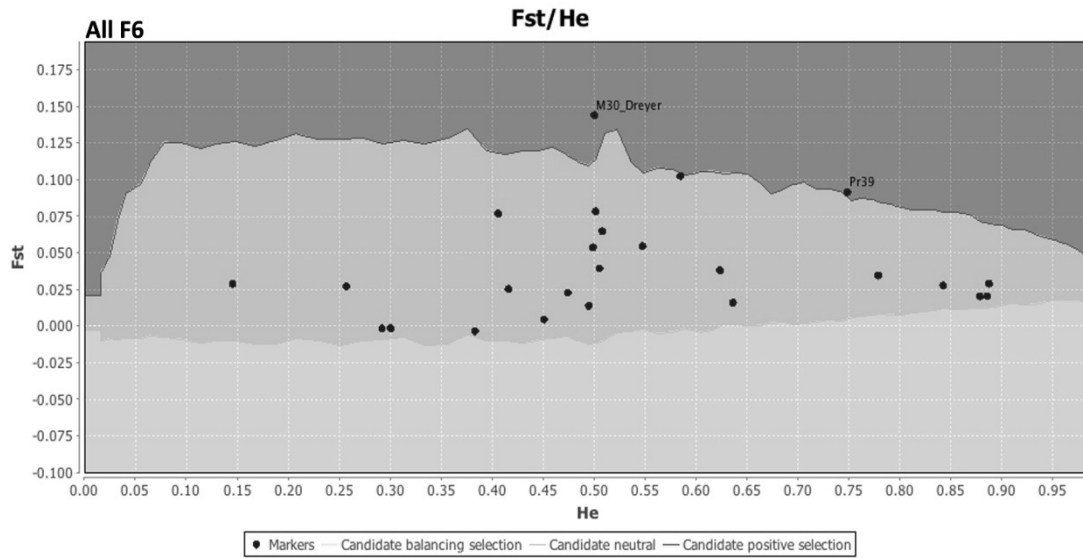
● Markers — Candidate balancing selection — Candidate neutral — Candidate positive selection

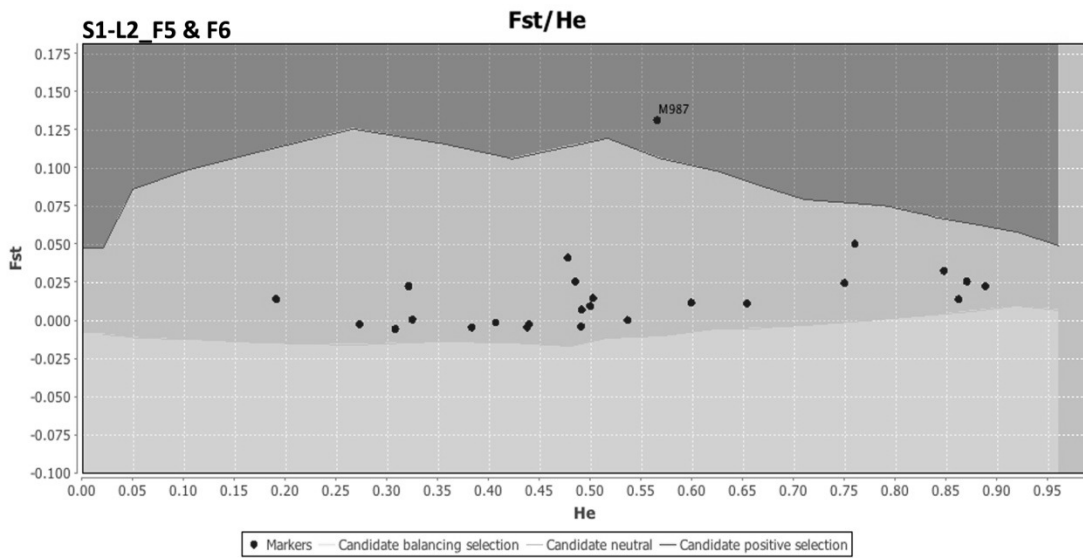
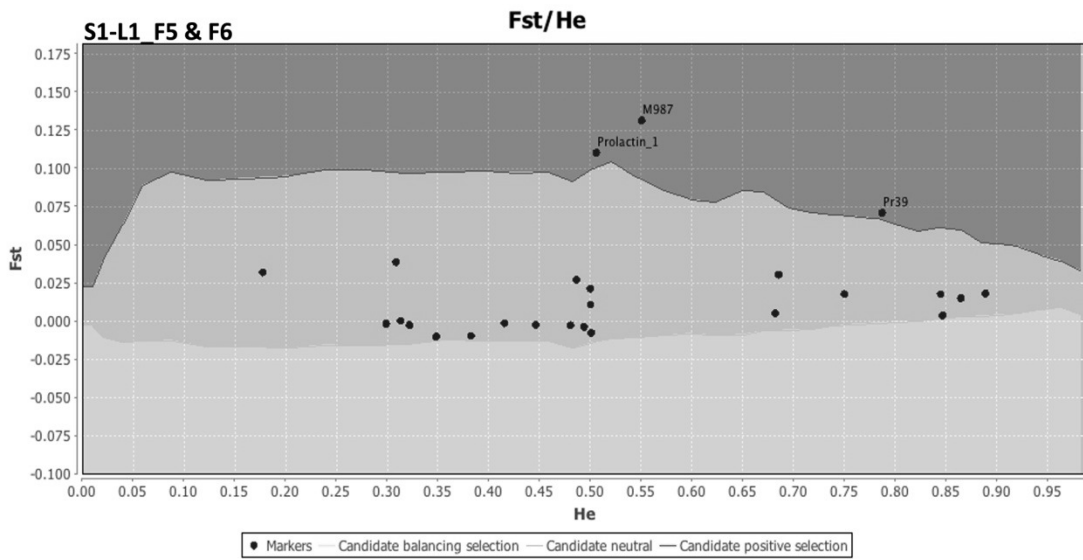
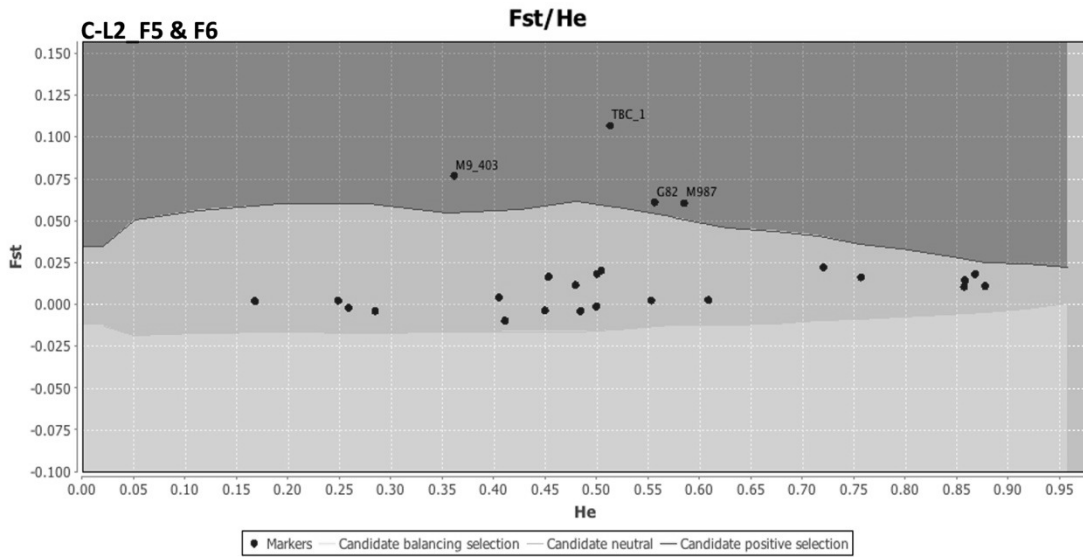


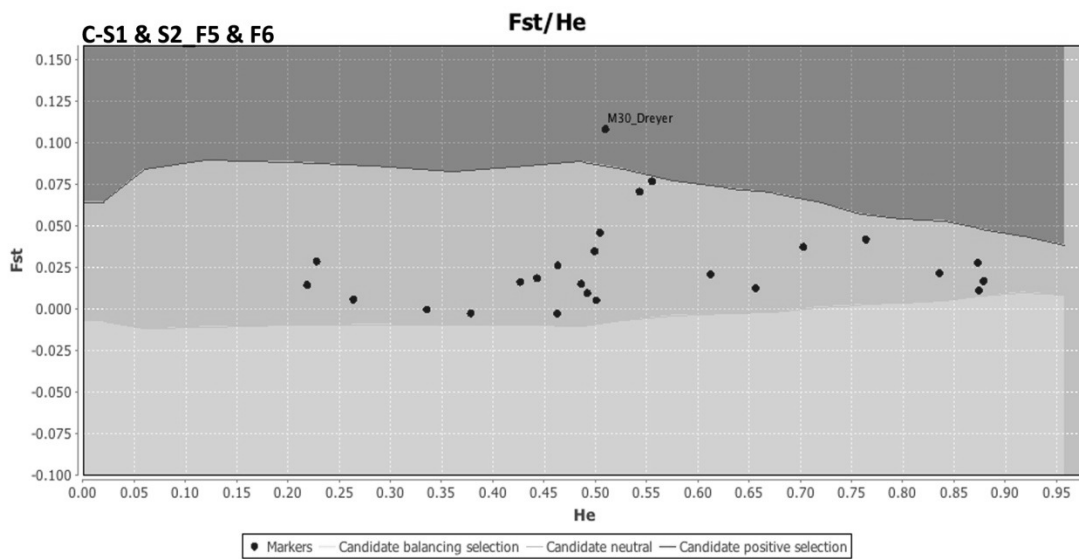
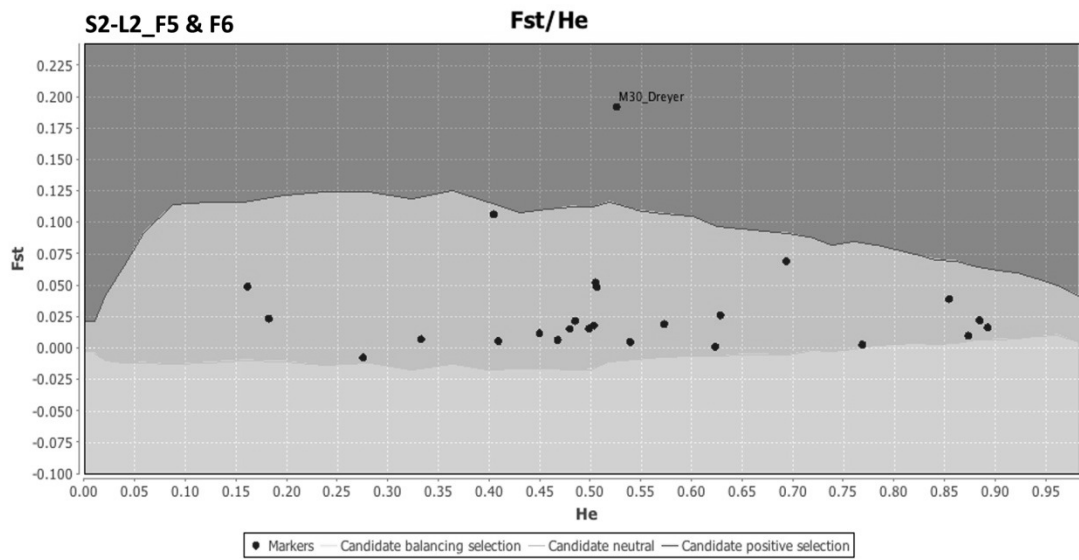
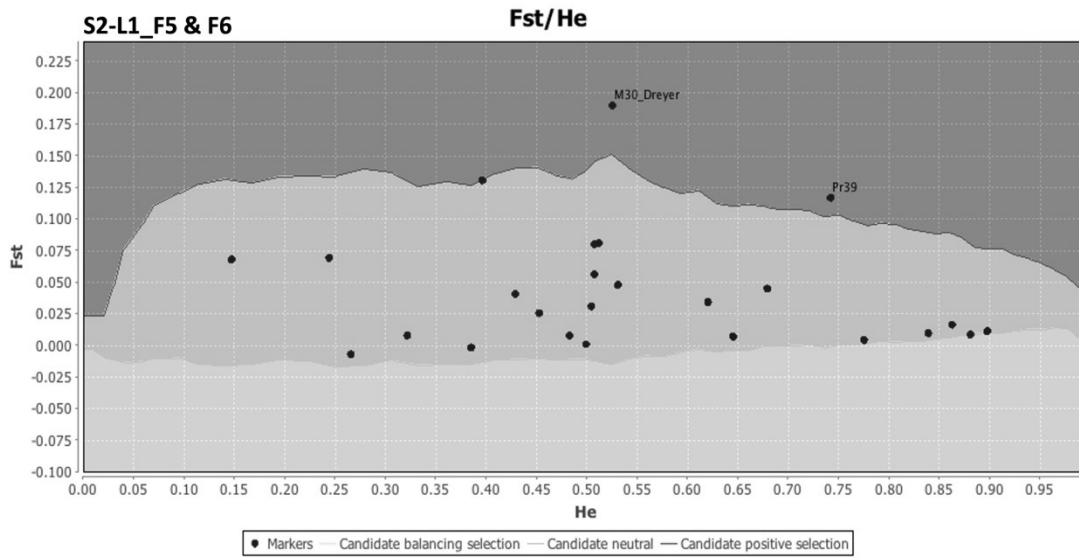
● Markers — Candidate balancing selection — Candidate neutral — Candidate positive selection



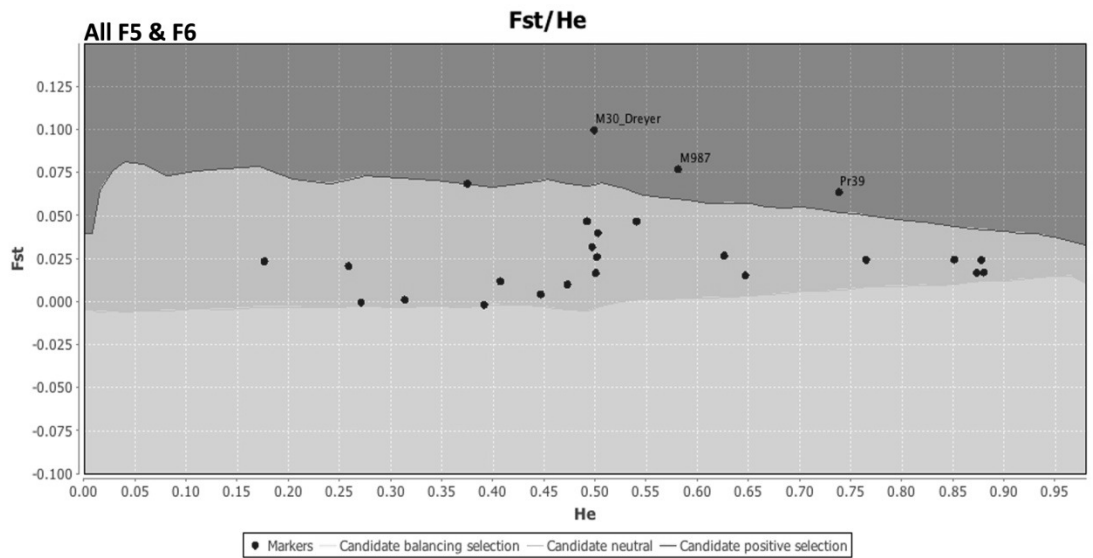
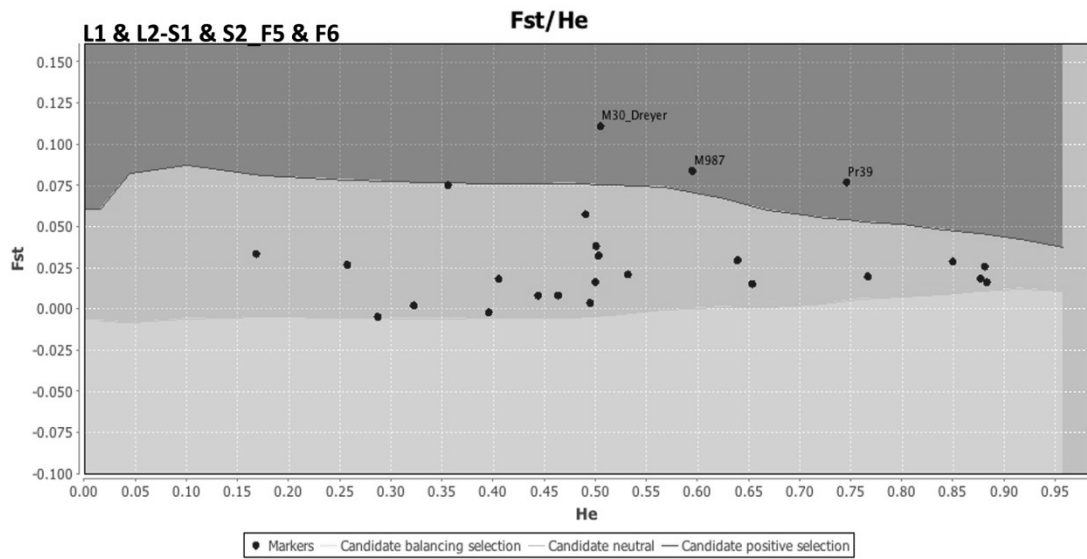
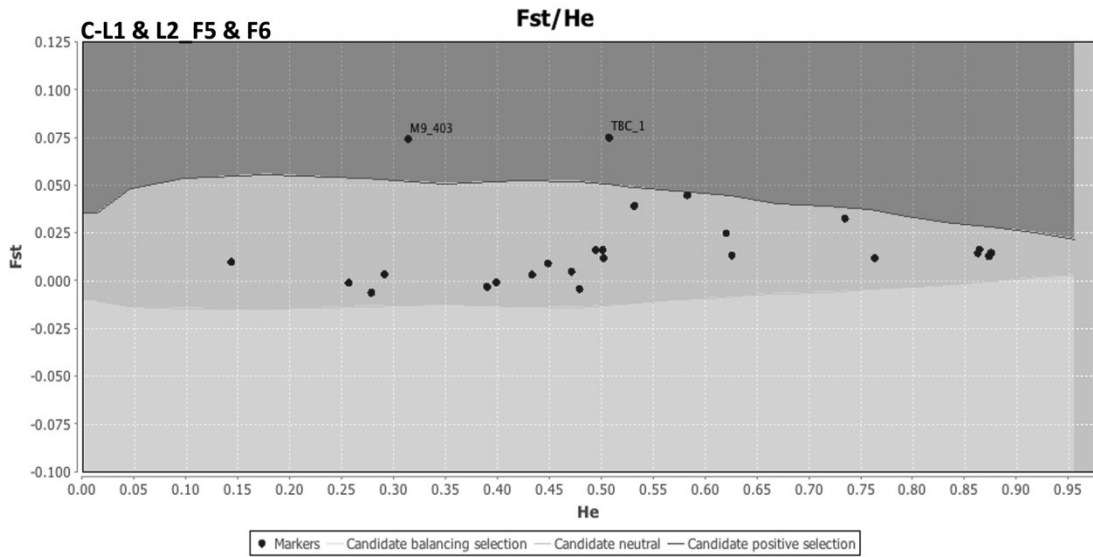












## APPENDIX VIII – Raw- and normalised ln(RH) data

- Expected heterozygosity at all microsatellite- and candidate loci for each line in generations F5 and F6. S-lines are shown in light grey, the Control in white and L-lines in dark grey.
- Raw ln(RH) values for each pairwise comparison within each generation. Ln(RH) is calculated as the natural logarithm of the ratio of heterozygosity of the population on the left of the pair over the population on the right of the pair:  $\ln(H_{e_{left}} / H_{e_{right}})$ .
- Normalised ln(RH) values so that, per population pair, the mean over all loci is 0 and the standard deviation equals 1.

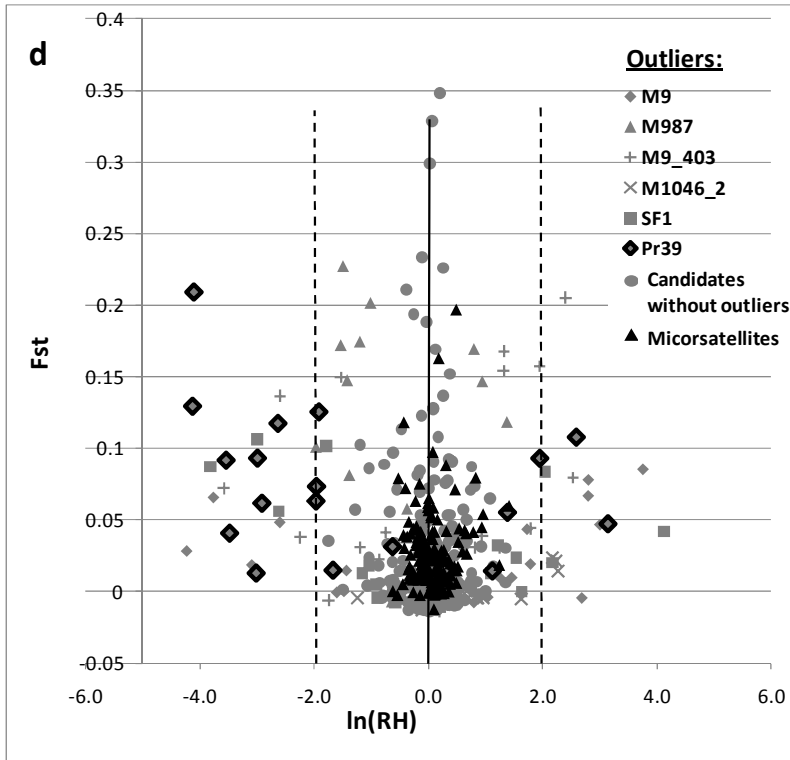
a.		Pret69	Pret-77	Hull 70-2	G102	Pret-32	G82	Pr92	G289	Pr39	M9_indel	M987
F5	S1	0.791	0.702	0.858	0.632	0.858	0.523	0.709	0.832	0.726	0.260	0.356
	S2	0.832	0.686	0.891	0.590	0.911	0.524	0.763	0.846	0.636	0.306	0.607
	C	0.869	0.620	0.839	0.577	0.818	0.457	0.719	0.838	0.706	0.208	0.557
	L1	0.871	0.594	0.886	0.699	0.835	0.503	0.769	0.850	0.732	0.142	0.590
F6	L2	0.845	0.637	0.841	0.596	0.858	0.558	0.744	0.891	0.709	0.098	0.621
	S1	0.808	0.677	0.851	0.660	0.824	0.519	0.689	0.869	0.751	0.243	0.412
	S2	0.771	0.599	0.878	0.581	0.891	0.520	0.776	0.842	0.535	0.059	0.645
	C	0.833	0.601	0.864	0.554	0.881	0.530	0.732	0.863	0.700	0.213	0.443
F6	L1	0.847	0.682	0.893	0.722	0.855	0.473	0.777	0.853	0.710	0.040	0.547
	L2	0.833	0.567	0.847	0.479	0.849	0.543	0.781	0.878	0.696	0.149	0.565

b.		Pret69	Pret-77	Hull 70-2	G102	Pret-32	G82	Pr92	G289	Pr39	M9_indel	M987
F5	S1-S2	-0.051	0.023	-0.037	0.068	-0.060	-0.002	-0.073	-0.017	0.132	-0.163	-0.534
	L1-L2	0.031	-0.070	0.052	0.160	-0.027	-0.104	0.033	-0.047	0.032	0.371	-0.051
	S1-C	-0.095	0.124	0.023	0.091	0.047	0.135	-0.014	-0.007	0.028	0.223	-0.448
	S2-C	-0.044	0.101	0.060	0.023	0.107	0.137	0.059	0.010	-0.104	0.386	0.086
	L1-C	0.002	-0.043	0.055	0.193	0.020	0.096	0.068	0.015	0.036	-0.382	0.058
	L2-C	-0.028	0.027	0.002	0.032	0.047	0.200	0.034	0.062	0.004	-0.753	0.109
	S1-L1	-0.097	0.168	-0.032	-0.102	0.027	0.039	-0.082	-0.022	-0.008	0.605	-0.505
	S1-L2	-0.067	0.098	0.021	0.059	-0.001	-0.065	-0.049	-0.069	0.024	0.976	-0.556
F6	S2-L1	-0.046	0.145	0.006	-0.170	0.087	0.041	-0.009	-0.005	-0.141	0.768	0.028
	S2-L2	-0.016	0.075	0.058	-0.009	0.060	-0.063	0.024	-0.052	-0.109	1.139	-0.023
	S1-S2	0.047	0.124	-0.031	0.127	-0.078	-0.001	-0.118	0.032	0.339	1.416	-0.448
	L1-L2	0.017	0.185	0.052	0.411	0.007	-0.139	-0.004	-0.029	0.020	-1.315	-0.032
	S1-C	-0.030	0.119	-0.015	0.175	-0.066	-0.020	-0.060	0.007	0.070	0.132	-0.073
	S2-C	-0.077	-0.005	0.016	0.048	0.011	-0.020	0.058	-0.024	-0.269	-1.284	0.376
	L1-C	0.017	0.126	0.033	0.264	-0.030	-0.115	0.059	-0.011	0.014	-1.672	0.211
	L2-C	0.000	-0.058	-0.019	-0.146	-0.037	0.024	0.064	0.018	-0.006	-0.357	0.243
F6	S1-L1	-0.047	-0.007	-0.048	-0.089	-0.036	0.095	-0.120	0.019	0.056	1.804	-0.283
	S1-L2	-0.031	0.177	0.004	0.322	-0.030	-0.044	-0.124	-0.010	0.076	0.489	-0.316
	S2-L1	-0.094	-0.131	-0.017	-0.216	0.041	0.095	-0.002	-0.013	-0.283	0.389	0.165
	S2-L2	-0.077	0.054	0.035	0.195	0.048	-0.043	-0.006	-0.042	-0.263	-0.926	0.132

c.		Pret69	Pret-77	Hull 70-2	G102	Pret-32	G82	Pr92	G289	Pr39	M9_indel	M987
F5	S1-S2	-0.216	0.068	-0.164	0.240	-0.253	-0.030	-0.301	-0.085	0.488	-0.646	-2.070
	L1-L2	0.373	-0.311	0.521	1.253	-0.021	-0.540	0.392	-0.154	0.382	2.681	-0.182
	S1-C	-0.627	0.614	0.040	0.424	0.175	0.675	-0.171	-0.128	0.068	1.173	-2.622
	S2-C	-0.181	0.302	0.166	0.041	0.321	0.422	0.160	-0.001	-0.382	1.247	0.250
	L1-C	0.159	-0.053	0.403	1.043	0.241	0.596	0.463	0.218	0.316	-1.626	0.416
	L2-C	-0.088	0.145	0.041	0.167	0.231	0.874	0.176	0.293	0.049	-3.131	0.489
	S1-L1	-0.553	0.458	-0.303	-0.570	-0.080	-0.035	-0.495	-0.265	-0.213	2.124	-2.108
	S1-L2	-0.337	0.277	-0.010	0.132	-0.090	-0.332	-0.270	-0.345	0.001	3.564	-2.170
F6	S2-L1	-0.219	0.253	-0.091	-0.524	0.111	-0.003	-0.127	-0.117	-0.452	1.793	-0.035
	S2-L2	-0.090	0.150	0.107	-0.072	0.111	-0.215	0.017	-0.186	-0.337	2.976	-0.108
	S1-S2	-0.087	0.108	-0.283	0.117	-0.401	-0.207	-0.504	-0.124	0.654	3.379	-1.340
	L1-L2	0.140	0.689	0.257	1.431	0.106	-0.370	0.070	-0.010	0.150	-4.226	-0.022
	S1-C	-0.398	0.447	-0.312	0.767	-0.603	-0.342	-0.568	-0.184	0.172	0.520	-0.637
	S2-C	-0.112	0.114	0.178	0.280	0.163	0.067	0.309	0.052	-0.713	-3.890	1.304
	L1-C	0.180	0.454	0.221	0.802	0.062	-0.151	0.287	0.110	0.173	-4.063	0.667
	L2-C	0.163	-0.166	0.053	-0.658	-0.044	0.295	0.520	0.261	0.130	-1.841	1.525
F6	S1-L1	-0.367	-0.263	-0.370	-0.475	-0.339	0.000	-0.555	-0.197	-0.100	4.422	-0.978
	S1-L2	-0.475	0.519	-0.309	1.209	-0.471	-0.541	-0.923	-0.378	0.034	2.010	-1.840
	S2-L1	-0.341	-0.457	-0.098	-0.727	0.087	0.258	-0.048	-0.086	-0.938	1.186	0.478
	S2-L2	-0.218	0.221	0.159	0.694	0.201	-0.105	0.020	-0.101	-0.843	-3.070	0.485



- d. Plot of normalised  $\ln(RH)$  values with Pr39 and M987 included as bi-allelic loci. The plot shows that, when only the presence/ absence of allele 174 at Pr39 is considered, the analysis does identify this locus as an outlier.



Pr39		
F5	S1-S2	-0.6266
F6	S1-S2	<b>2.5919</b>
F5	L1-L2	0.2535
F6	L1-L2	0.2785
F5	S1-C	<b>-3.0141</b>
F6	S1-C	<b>-1.9677</b>
F5	S2-C	-1.6694
F6	S2-C	<b>-3.4803</b>
F5	L1-C	1.3859
F6	L1-C	1.9510
F5	L2-C	1.1207
F6	L2-C	<b>3.1453</b>
F5	S1-L1	<b>-2.9910</b>
F6	S1-L1	<b>-2.6341</b>
F5	S1-L2	<b>-2.9145</b>
F6	S1-L2	<b>-3.5415</b>
F5	S2-L1	-1.9167
F6	S2-L1	<b>-4.1038</b>
F5	S2-L2	-1.9578
F6	S2-L2	<b>-4.1268</b>

## APPENDIX IX – Association analyses models of inheritance

Models of inheritance for all candidate markers showing a significant association between standard length and genotype in at least one of the four test groups: 1) over the F0, F2, all populations in the F5 and the F6 combined, 2), within the F5, 3) within the F6 and 4) over all of the Control line populations (F0, F2, F5\_C and F6\_C). Shown are sample size (n), mean response associated with the respective genotype (R), standard error of the mean and the difference between the genotypes in mm. For significant associations, furthermore are shown the 95% confidence intervals of the difference and the *p*-values associated with the test. The test value is shown in bold and the *p*-value after Benjamini-Hochberg correction is shown below each *p*-value. AIC in the final column stands for the Akaike Information Criterion. The AIC is useful to decide on the most likely model of inheritance, which should be supported by the lowest AIC. The most likely model of inheritance for each test is shown in grey.

### F0, F2, All F5, All F6

<u>M9_indel</u>	Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	no insert	498	19.72	0.09	0.00			
	insert/ no insert	133	19.11	0.18	<b>-0.62</b>	( -1.01 - -0.23 )	<b>0.0071</b>	2720.4
	2 copies							
	insert	4	20.3	0.56	0.57		0.0196	

### All F5

<u>M9_indel</u>	Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	no insert	146	20.32	0.23	0.00			
	insert/ no insert	44	18.54	0.41	<b>-1.78</b>	( -2.69 - -0.86 )	<b>0.0002</b>	923.8
	2 copies							
	insert	---					0.0012	

### All F6

<u>M9_indel</u>	Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	no insert	207	19.21	0.11	0.00			
	insert/ no insert	37	19.05	0.25	-0.16		n.s.	
	2 copies							
	insert	---						

### Control only

<u>M9_indel</u>	Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	no insert	213	19.72	0.11	0.00			
	insert/ no insert	73	19.51	0.19	-0.21		n.s.	
	2 copies							
	insert	---						

### F0, F2, All F5, All F6

<u>SF1</u>	Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Overdominant	a/a-t/t	466	19.52	0.1	0.00		n.s.	
	a/t	168	19.86	0.16	0.34			

### All F5

<u>SF1</u>	Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Overdominant	a/a-t/t	137	19.58	0.24	0.00		<b>0.0092</b>	918.1
	a/t	50	20.79	0.37	<b>1.21</b>	( 0.31 - 2.12 )	0.0339	

### All F6

<u>SF1</u>	Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Overdominant	a/a-t/t	178	19.21	0.11	0.00		n.s.	
	a/t	64	19.17	0.22	-0.04			

### Control only

<u>SF1</u>	Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Overdominant	a/a-t/t	214	19.67	0.11	0.00		n.s.	
	a/t	78	19.71	0.17	0.04			

F0, F2, All F5, All F6

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	334	19.83	0.12	0.00	( -0.71 - -0.03 )	<b>0.0032</b>	2677.1
	241	19.46	0.13	<b>-0.04</b>	( -1.55 - -0.34 )	0.0109	
	51	18.89	0.22	<b>-0.94</b>			
Dominant	334	19.83	0.12	0.00	( -0.79 - -0.15 )	<b>0.0042</b>	2678.4
	292	19.36	0.11	<b>-0.47</b>		0.0132	
Recessive	575	19.68	0.09	0.00	( -1.37 - -0.20 )	<b>0.0088</b>	2679.8
	51	18.89	0.22	<b>-0.79</b>		0.0213	
Overdominant	385	19.71	0.11			n.s.	
	241	19.46	0.13				
Log additive	---	---	---	<b>-0.43</b>	( -0.68 - -0.18 )	<b>0.0008</b>	2675.4
	---	---	---			0.0036	

All F5

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	117	20.54	0.26	0.00	( -2.31 - -0.58 )	<b>0.0006</b>	914.7
	56	19.09	0.34	<b>-1.45</b>	( -3.55 - -0.63 )	0.0029	
	15	18.45	0.55	<b>-2.09</b>			
Dominant	117	20.54	0.26	0.00	( -2.38 - -0.78 )	<b>0.0001</b>	913.4
	71	18.95	0.29	<b>-1.58</b>		0.0006	
Recessive	173	20.07	0.22	0.00	( -3.09 - -0.15 )	<b>0.0320</b>	923.3
	15	18.45	0.55	<b>-1.62</b>		n.s.	
Overdominant	132	20.3	0.25	0.00	( -2.08 - -0.35 )	<b>0.0067</b>	920.5
	56	19.09	0.34	<b>-1.21</b>		0.0271	
Log additive	---	---	---	<b>-1.20</b>	( -1.81 - -0.60 )	<b>0.0001</b>	913.3
	---	---	---			0.0006	

Control only

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	126	19.39	0.14	0.00		n.s.	
	96	19.01	0.16	-0.38			
	20	18.91	0.21	-0.48			
Dominant	126	19.39	0.14	0.00		n.s.	
	116	18.99	0.14	-0.39			
Recessive	222	19.22	0.11	0.00		n.s.	
	20	18.91	0.21	-0.32			
Overdominant	146	19.32	0.13	0.00		n.s.	
	96	19.01	0.16	-0.31			
Log additive	---	---	---	-0.30		n.s.	

All F6

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	130	19.57	0.13	0.00			
	128	19.89	0.14	0.32		<b>0.032</b>	
	27	19.08	0.27	-0.49	( -1.14 - 0.16 )	n.s.	
Dominant	130	19.57	0.13	0.00		n.s.	
	155	19.75	0.13	0.18			
Recessive	258	19.73	0.1	0.00	( -1.27 - -0.03 )	<b>0.041</b>	1069.4
	27	19.08	0.27	<b>-0.65</b>		n.s.	
Overdominant	157	19.49	0.12	0.00	( 0.04 - 0.77 )	<b>0.031</b>	1068.8
	128	19.89	0.14	<b>0.41</b>		n.s.	
Log additive	---	---	---	-0.03		n.s.	

FO, F2, All F5, All F6

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	c/c	271	19.88	0.13	0.00		
	c/t	299	19.37	0.11	<b>-0.52</b>	( -0.85 - -0.18 )	<b>0.0094</b>
	t/t	56	19.44	0.3	-0.45		0.0216
Dominant	c/c	271	19.88	0.13	0.00		2675.3
	c/c-t/t	355	19.38	0.11	<b>-0.50</b>	( -0.83 - -0.18 )	<b>0.0081</b>
Overdominant	c/c-t/t	327	19.81	0.12	0.00		2677.5
	c/t	299	19.37	0.11	<b>-0.44</b>	( -0.76 - -0.12 )	0.0196
Log additive	---	---	---	---	<b>-0.34</b>	( -0.59 - -0.09 )	<b>0.0080</b>
	---	---	---	---		0.0200	2677.6

All F5

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	c/c	82	20.15	0.32	0.00		
	c/t	83	19.6	0.29	-0.55		n.s.
	t/t	19	19.86	0.72	-0.29		
Dominant	c/c	82	20.15	0.32	0.00		n.s.
	c/c-t/t	102	19.65	0.27	-0.50		
Overdominant	c/c-t/t	101	20.1	0.29	0.00		n.s.
	c/t	83	19.6	0.29	-0.50		
Log additive	---	---	---	---		n.s.	

All F6

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	c/c	100	19.3	0.15	0.00		
	c/t	118	118	0.15	-0.13		n.s
	t/t	24	24	0.34	-0.54		
Dominant	c/c	100	100	0.15	0.00		n.s
	c/c-t/t	142	142	0.14	-0.20		
Overdominant	c/c-t/t	124	124	0.14	0.00		n.s
	c/t	118	118	0.15	-0.03		
Log additive	---	---	---	---	<b>-0.22</b>		n.s

Control only

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	c/c	126	20.03	0.14	0.00		
	c/t	141	19.41	0.14	<b>-0.62</b>	( -1.00 - -0.24 )	<b>0.0065</b>
	t/t	22	19.64	0.29	-0.38		<b>0.03934</b>
Dominant	c/c	126	20.03	0.14	0.00		1086.9
	c/c-t/t	163	19.44	0.13	<b>-0.59</b>	( -0.95 - -0.22 )	<b>0.02314</b>
Overdominant	c/c-t/t	148	19.97	0.12	0.00		1087.6
	c/t	141	19.41	0.14	<b>-0.56</b>	( -0.92 - -0.20 )	<b>0.0243</b>
Log additive	---	---	---	---		<b>0.0110</b>	1090
	---	---	---	---		0.0557	

F0, F2, All F5, All F6

GH2\_60

Model	n		R	St. err.	Difference	95% CI	p-value	AIC
	n	R						
Codominant	g/g	151	20.14	0.15	0.00			
	g/a	333	19.54	0.11	<b>-0.60</b>	( -0.99 - -0.21 )	<b>0.0004</b>	2669.1
	a/a	142	19.22	0.17	<b>-0.92</b>	( -1.39 - -0.46 )	0.0021	
Dominant	g/g	151	20.14	0.15	0.00		<b>0.0003</b>	2669.7
	g/a-a/a	475	19.45	0.1	<b>-0.70</b>	( -1.07 - -0.32 )	0.0017	
Recessive	g/g-g/a	484	19.73	0.09	0.00		<b>0.0090</b>	2676.1
	a/a	142	19.22	0.17	<b>-0.51</b>	( -0.89 - -0.13 )	0.0213	
Log additive	---	---	---	---	<b>-0.46</b>	( -0.70 - -0.23 )	<b>0.0001</b>	2667.8
							0.0007	

All F5

GH2\_60

Model	n		R	St. err.	Difference	95% CI	p-value	AIC
	n	R						
Codominant	g/g	29	21.19	0.52	0.00			
	g/a	110	19.75	0.26	<b>-1.44</b>		<b>0.0290</b>	895.5
	a/a	44	19.58	0.41	<b>-1.61</b>	( -2.90 - -0.32 )	n.s.	
Dominant	g/g	29	21.19	0.52	0.00		<b>0.0083</b>	893.7
	g/a-a/a	154	19.7	0.22	<b>-1.49</b>	( -2.58 - -0.39 )	0.0320	
Recessive	g/g-g/a	139	20.05	0.24	0.00		n.s.	
	a/a	44	19.58	0.41	<b>-0.47</b>		<b>0.0280</b>	895.8
Log additive	---	---	---	---	<b>-0.73</b>	( 0.09 - 1.37 )	n.s.	

All F6

GH2\_60

Model	n		R	St. err.	Difference	95% CI	p-value	AIC
	n	R						
Codominant	g/g	64	19.54	0.18	0.00			
	g/a	119	19.11	0.15	-0.43		n.s.	
	a/a	59	19	0.21	-0.54			
Dominant	g/g	64	19.54	0.18	0.00		<b>0.043</b>	906.7
	g/a-a/a	178	19.07	0.12	<b>-0.46</b>	( -0.91 - -0.02 )	n.s.	
Recessive	g/g-g/a	183	19.26	0.11	0.00		n.s.	
	a/a	59	19	0.21	-0.26			
Log additive	---	---	---	---	<b>-0.27</b>		n.s.	

Control only

GH2\_60

Model	n		R	St. err.	Difference	95% CI	p-value	AIC
	n	R						
Codominant	g/g	71	20.14	0.18	0.00			
	g/a	155	19.61	0.12	<b>-0.53</b>	( -0.98 - -0.09 )	<b>0.015</b>	1094.3
	a/a	63	19.38	0.23	<b>-0.76</b>	( -1.30 - -0.22 )	n.s.	
Dominant	g/g	71	20.14	0.18	0.00		<b>0.0062</b>	1093.3
	g/a-a/a	218	19.54	0.11	<b>-0.60</b>	( -1.03 - -0.17 )	0.03934	
Recessive	g/g-g/a	226	19.77	0.1	0.00		n.s.	
	a/a	63	19.38	0.23	-0.40			
Log additive	---	---	---	---	<b>-0.39</b>	( -0.66 - -0.12 )	<b>0.0053</b>	1093
							0.0390	



F0, F2, All F5, All F6

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
t/t	192	19.88	0.15	0.00			
Codominant t/c	309	19.81	0.12	-0.07		<0.0001	2681.6
c/c	129	18.76	0.15	-1.12	( -1.58 - -0.67 )	0.0008	
t/t	192	19.88	0.15	0.00		0.0340	2704
Dominant t/c-c/c	438	19.5	0.1	-0.37	( -0.73 - -0.03 )	n.s.	
t/t-t/c	501	19.84	0.09	0.00		<0.0001	2679.8
c/c	129	18.76	0.15	-1.08	( -1.47 - -0.69 )	0.0008	
Overdominant t/t-c/c	321	19.43	0.11	0.00		0.0200	2703.1
t/c	309	19.81	0.12	0.38	( 0.06 - 0.71 )	0.0405	
Log additive ---	---	---	---	-0.51	( -0.74 - -0.29 )	<0.0001	2688.8
						0.0008	

All F5

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
t/t	48	20.95	0.41	0.00			
Codominant t/c	100	20.24	0.27	-0.72		<0.0001	894.4
c/c	39	17.89	0.33	-3.07	( -4.17 - -1.96 )	0.0001	
t/t	48	20.95	0.41	0.00		0.0033	914.1
Dominant t/c-c/c	139	19.58	0.23	-1.38	( -2.28 - -0.47 )	0.0141	
t/t-t/c	148	20.47	0.23	0.00		<0.0001	894.8
c/c	39	17.89	0.33	-2.58	( -3.51 - -1.66 )	0.0001	
Overdominant t/t-c/c	87	19.58	0.31	0.00		n.s.	
t/c	100	20.24	0.27	0.66		<0.0001	896.9
Log additive ---	---	---	---	-1.49	( -2.04 - -0.93 )	0.0001	

All F6

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
t/t	86	19.28	0.17	0.00			
Codominant t/c	109	19.35	0.17	0.07		0.028	902.9
c/c	46	18.64	0.16	-0.65	( -1.20 - -0.09 )	n.s.	
t/t	86	19.28	0.17	0.00		n.s.	
Dominant t/c-c/c	155	19.14	0.13	-0.14			
t/t-t/c	195	19.32	0.12	0.00		0.0079	901
c/c	46	18.64	0.16	-0.68	( -1.18 - -0.18 )	n.s.	
Overdominant t/t-c/c	132	19.06	0.13	0.00		n.s.	
t/c	109	19.35	0.17	0.29			
Log additive ---	---	---	---	-0.27		n.s.	

Control only

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
t/t	93	19.57	0.17	0.00			
Codominant t/c	146	19.83	0.14	0.26		n.s.	
c/c	52	19.55	0.18	-0.02			
t/t	93	19.57	0.17	0.00		n.s.	
Dominant t/c-c/c	198	19.75	0.11	0.18			
t/t-t/c	239	19.73	0.11	0.00		n.s.	
c/c	52	19.55	0.18	-0.18			
Overdominant t/t-c/c	145	19.56	0.13	0.00		n.s.	
t/c	146	19.83	0.14	0.26			
Log additive ---	---	---	---	0.03		n.s.	

**FO, F2, All F5, All F6**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
a/a	367	19.43	0.11	0.00			
Codominant	195	19.86	0.15	<b>0.43</b>	( 0.07 - 0.79 )	<b>0.0170</b>	2683.6
c/c	64	20.02	0.24	<b>0.59</b>	( 0.05 - 1.14 )	0.0361	
Dominant	367	19.43	0.11	0.00		<b>0.0049</b>	2681.9
c/a-c/c	259	19.9	0.13	<b>0.17</b>	( 0.14 - 0.80 )	0.0149	
Log additive	---	---	---	<b>0.34</b>	( 0.10 - 0.58 )	<b>0.0053</b>	2682.1
	---	---	---	---		0.0155	

**All F5**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
a/a	102	19.56	0.27	0.00			
Codominant	67	20.24	0.36	0.68		n.s.	
c/c	19	20.85	0.63	1.30			
Dominant	102	19.56	0.27	0.00		<b>0.0470</b>	924.8
c/a-c/c	86	20.38	0.31	<b>0.82</b>	( 0.02 - 1.62 )	n.s.	
Log additive	---	---	---	<b>0.66</b>	( 0.06 - 1.26 )	<b>0.0310</b>	924
	---	---	---	---		n.s.	

**Myostatin**

**All F6**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
a/a	140	19.04	0.14	0.00			
Codominant	77	19.45	0.18	0.41		n.s.	
c/c	24	19.29	0.27	0.25			
Dominant	140	19.04	0.14	0.00		n.s.	
c/a-c/c	101	19.41	0.15	0.37			
Log additive	---	---	---	0.22		n.s.	

**Control only**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
a/a	178	19.63	0.12	0.00			
Codominant	80	19.77	0.17	0.14		n.s.	
c/c	28	20.04	0.25	0.42			
Dominant	178	19.63	0.12	0.00		n.s.	
c/a-c/c	108	19.84	0.14	0.21			
Log additive	---	---	---	0.19		n.s.	

**Myostatin**

**FO, F2, All F5, All F6**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
g/g	229	19.26	0.13	0.00			
codominant	296	19.66	0.12	<b>0.40</b>	( 0.05 - 0.76 )	<b>0.0003</b>	2711.8
a/a	109	20.22	0.2	<b>0.96</b>	( 0.50 - 1.43 )	0.0017	
Dominant	229	19.26	0.13	0.00		<b>0.0012</b>	2715.8
g/a-a/a	405	19.81	0.1	<b>0.55</b>	( 0.22 - 0.89 )	0.0049	
g/g-g/a	525	19.48	0.09	0.00		<b>0.0007</b>	2714.8
a/a	109	20.22	0.2	<b>0.74</b>	( 0.31 - 1.16 )	0.0033	
Log additive	---	---	---	<b>0.47</b>	( 0.24 - 0.69 )	<b>0.0001</b>	2710
	---	---	---	---		0.0007	

**All F5**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
g/g	61	19.25	0.34	0.00			
codominant	100	19.87	0.28	0.61		<b>0.0009</b>	916.7
a/a	27	21.65	0.48	<b>2.40</b>	( 1.16 - 3.64 )	0.0041	
Dominant	61	19.25	0.34	0.00		<b>0.0240</b>	923.7
g/a-a/a	127	20.25	0.25	<b>0.99</b>	( 0.14 - 1.85 )	n.s.	
g/g-g/a	161	19.64	0.22	0.00		<b>0.0005</b>	916.7
a/a	27	21.65	0.48	<b>2.01</b>	( 0.90 - 3.13 )	0.0025	
Log additive	---	---	---	<b>1.07</b>	( 0.48 - 1.66 )	<b>0.0005</b>	916.7
	---	---	---	---		0.0025	

**Prolactin1**

**All F6**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
g/g	90	18.93	0.17	0.00			
codominant	97	19.3	0.15	0.37		n.s.	
a/a	56	19.41	0.21	0.47			
Dominant	90	18.93	0.17	0.00		<b>0.048</b>	911.1
g/a-a/a	153	19.34	0.12	<b>0.41</b>	( 0.01 - 0.82 )	n.s.	
g/g-g/a	187	19.12	0.11	0.00		n.s.	
a/a	56	19.41	0.21	0.29			
Log additive	---	---	---	0.25		n.s.	

**Control only**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
g/g	101	19.68	0.16	0.00			
codominant	146	19.63	0.13	-0.06		n.s.	
a/a	43	19.88	0.26	0.20			
Dominant	101	19.68	0.16	0.00		n.s.	
g/a-a/a	189	19.69	0.12	0.00			
g/g-g/a	247	19.65	0.1	0.00		n.s.	
a/a	43	19.88	0.26	0.23			
Log additive	---	---	---	0.07		n.s.	

**Prolactin1**

**F0, F2, All F5, All F6**  
**Prolactin2**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	185	20.05	0.16	0.00	( -0.89 - -0.13 )	<b>0.0021</b>	2687.6
	280	19.54	0.12	<b>-0.51</b>	( -1.18 - -0.32 )	0.0078	
Dominant	185	20.05	0.16	0.00		<b>0.0010</b>	2687
	442	19.45	0.1	<b>-0.60</b>	( -0.95 - -0.24 )	0.0043	2692.4
Recessive	465	19.74	0.1	0.00		<b>0.0190</b>	
	162	19.3	0.16	-0.44	( -0.81 - -0.08 )	0.0394	
Overdominant	347	19.7	0.11	0.00		n.s.	
	280	19.54	0.12	<b>-0.16</b>		<b>0.0006</b>	2686.2
Log additive	---	---	---	<b>-0.38</b>	( -0.60 - -0.16 )	0.0030	

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	48	20.37	0.43	0.00		n.s.	
	85	20.08	0.31	-0.29			
Dominant	52	19.4	0.37	-0.97		n.s.	
	48	20.37	0.43	0.00			
Recessive	137	19.82	0.24	-0.55			
	133	20.18	0.25	0.00		n.s.	
Overdominant	52	19.4	0.37	-0.78			
	100	19.87	0.28	0.00		n.s.	
Log additive	85	20.08	0.31	0.21		n.s.	
	---	---	---	<b>-0.49</b>			

**Control only**  
**Prolactin2**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	84	19.36	0.17	0.00		n.s.	
	99	19.18	0.15	-0.19			
Dominant	56	18.97	0.24	-0.40		n.s.	
	84	19.36	0.17	0.00			
Recessive	155	19.1	0.13	-0.26		n.s.	
	183	19.26	0.11	0.00			
Overdominant	56	18.97	0.24	-0.30		n.s.	
	140	19.2	0.14	0.00			
Log additive	99	19.18	0.15	-0.03		n.s.	
	---	---	---	<b>-0.20</b>			

**All F6**  
**Prolactin2**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	75	19.55	0.18	0.00		<b>&lt;0.0001</b>	1077.5
	137	19.39	0.12	-0.17	( 0.33 - 1.33 )	0.00041	
Dominant	76	20.38	0.2	<b>0.83</b>		<b>&lt;0.0001</b>	1076
	76	20.38	0.2	0.00	( -1.35 - -0.53 )	0.00041	
Recessive	212	19.45	0.1	<b>-0.94</b>		n.s.	
	213	19.74	0.11	0.00			
Overdominant	75	19.55	0.18	-0.19		<b>0.002</b>	1086.1
	151	19.97	0.14	0.00	( -0.95 - -0.22 )	0.02314	
Log additive	137	19.39	0.12	<b>-0.58</b>	( -0.67 - -0.16 )	<b>0.0013</b>	1085.4
	---	---	---	<b>-0.42</b>		0.0231	

**F0, F2, All F5, All F6**  
**IBC.1**

**All F5**  
**IBC.1**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	g/g	165	19.72	0.16	0.00		
	t/g	317	19.74	0.12	0.02	<b>0.0480</b>	2714.9
	t/t	151	19.26	0.15	<b>-0.46</b>	( -0.91 - -0.01 )	n.s.
Dominant	g/g	165	19.72	0.16	0.00	n.s.	
	t/g-t/t	468	19.58	0.1	-0.13		
Recessive	g/g-t/g	482	19.73	0.1	0.00	<b>0.0140</b>	2712.9
	t/t	151	19.26	0.15	<b>-0.47</b>	( -0.85 - -0.10 )	0.0305
Overdominant	g/g-t/t	316	19.5	0.11	0.00	n.s.	
	t/g	317	19.74	0.12	0.24		
Log additive	---	---	---	---	-0.22	n.s.	

**All F6**  
**IBC.1**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	g/g	67	19.73	0.16	0.00		
	t/g	110	19.1	0.15	<b>-0.63</b>	( -1.10 - -0.16 )	<b>0.0016</b>
	t/t	66	18.78	0.2	<b>-0.94</b>	( -1.47 - -0.42 )	0.0154
Dominant	g/g	67	19.73	0.16	0.00	<b>0.0008</b>	903.6
	t/g-t/t	176	18.98	0.12	<b>-0.75</b>	( -1.18 - -0.32 )	0.0088
Recessive	g/g-t/g	177	19.34	0.12	0.00	<b>0.015</b>	908.9
	t/t	66	18.78	0.2	<b>-0.55</b>	( -0.99 - -0.11 )	n.s.
Overdominant	g/g-t/t	133	19.26	0.13	0.00	n.s.	
	t/g	110	19.1	0.15	-0.16		
Log additive	---	---	---	---	<b>-0.47</b>	<b>0.0005</b>	902.6
					( -0.73 - -0.21 )	0.0064	

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	g/g	45	19.83	0.46	0.00		
	t/g	102	20.18	0.28	0.35	n.s.	
	t/t	40	19.53	0.38	-0.31		
Dominant	g/g	45	19.83	0.46	0.00	n.s.	
	t/g-t/t	142	20	0.23	0.17		
Recessive	g/g-t/g	147	20.08	0.23	0.00	n.s.	
	t/t	40	19.53	0.38	-0.55		
Overdominant	g/g-t/t	85	19.69	0.3	0.00	n.s.	
	t/g	102	20.18	0.28	0.50		
Log additive	---	---	---	---	-0.14	n.s.	

**Control only**  
**IBC.1**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	t/t	87	19.33	0.16	0.00		
	t/g	141	19.97	0.14	<b>0.64</b>	( 0.21 - 1.06 )	<b>0.011</b>
	g/g	63	19.56	0.2	0.23	n.s.	1099.5
Dominant	g/g	63	19.56	0.2	0.00	n.s.	
	t/g-t/t	228	19.72	0.11	0.17		
Recessive	g/g-t/g	204	19.84	0.11	0.00	<b>0.013</b>	1100.4
	t/t	87	19.33	0.16	<b>-0.51</b>	( -0.91 - -0.11 )	n.s.
Overdominant	g/g-t/t	150	19.42	0.12	0.00	<b>0.0039</b>	1098.2
	t/g	141	19.97	0.14	<b>0.54</b>	( 0.18 - 0.91 )	0.03159
Log additive	---	---	---	---	<b>-0.15</b>	n.s.	

**F0, F2, All F5, All F6**

**Pr39**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
x/x	422	19.12	0.09	0.00			
Codominant	215	20.56	0.1	<b>1.44</b>	( 1.12 - 1.76 )	<0.0001	2673.7
174/174	2	19.27	3.13	0.15		0.0001	
Dominant	422	19.12	0.09	0.00		<0.0001	
174/x-	217	20.55	0.15	<b>1.43</b>	( 1.11 - 1.75 )	0.0001	2672.5
174/174						<0.0001	2675.4
Log additive	---	---	---	<b>1.38</b>	( 1.06 - 1.70 )	0.0001	

**All F5**

**Pr39**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
x/x	100	18.64	0.24	0.00			
Codominant	88	21.36	0.27	<b>2.72</b>	( 2.01 - 3.43 )	<0.0001	889.9
174/174	2	19.27	3.13	0.63		0.0001	
Dominant	100	18.64	0.24	0.00		<0.0001	889.3
174/x-	90	21.31	0.27	<b>2.67</b>	( 1.96 - 3.38 )	0.0001	
174/174						<0.0001	894.6
Log additive	---	---	---	<b>2.43</b>	( 1.75 - 3.12 )	0.0001	

**All F6**

**Pr39**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
x/x	146	18.63	0.1	0.00			
Codominant	98	20.01	0.17	<b>1.39</b>	( 1.02 - 1.75 )	<0.0001	867
174/174	---					0.000154	
Dominant							
Log additive	---						

**Control only**

**Pr39**

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
x/x	236	19.6	0.1	0.00			
Codominant	58	20.03	0.24	0.43		0.069	1113.2
174/174	---					n.s.	
Dominant							
Log additive	---						

F0, F2, All F5, All F6  
M98Z

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	258 317	18.91 20.04	0.11 0.12	0.00 1.13	( 0.80 - 1.45 ) ( 0.82 - 1.94 )	<0.0001 0.0001	2677.5
Dominant	258 376	18.91 20.08	0.11 0.11	0.00 1.17	( 0.85 - 1.48 )	<0.0001 0.0001	2676.3
Recessive	575 59	19.53 20.29	0.09 0.25	0.00 0.76	( 0.21 - 1.31 )	0.0074 0.0196	2719.7
Overdominant	317 317	19.17 20.04	0.11 0.12	0.00 0.87	( 0.55 - 1.19 ) ( 0.61 - 1.11 )	<0.0001 0.0001	2698.2 2681.3
Log additive	---	---	---	0.86		0.0001	

All F5  
M98Z

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	64 112	18.38 20.67	0.29 0.26	0.00 2.29	( 1.49 - 3.10 ) ( 0.89 - 3.91 )	<0.0001 0.0001	908.8
Dominant	64 126	18.38 20.69	0.29 0.24	0.00 2.31	( 1.52 - 3.09 )	<0.0001 0.0001	906.8
Recessive	176 14	19.84 20.78	0.21 0.69	0.00 0.94		n.s.	
Overdominant	78 112	18.81 20.67	0.28 0.26	0.00 1.86	( 1.09 - 2.63 ) ( 1.06 - 2.34 )	<0.0001 0.0001	916.5 912.5
Log additive	---	---	---	1.70		0.0001	

All F6  
M98Z

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	81 140	18.38 19.6	0.14 0.13	0.00 1.23	( 0.82 - 1.63 ) ( 0.43 - 1.79 )	<0.0001 0.000154	885.2
Dominant	81 163	18.38 19.59	0.14 0.12	0.00 1.21	( 0.82 - 1.60 )	<0.0001 0.000154	883.3
Recessive	221 23	19.15 19.49	0.1 0.35	0.00 0.33		n.s.	
Overdominant	104 140	18.62 19.6	0.14 0.13	0.00 0.98	( 0.50 - 1.36 ) ( 0.49 - 1.11 )	<0.0001 0.000154	893.3 893
Log additive	---	---	---	0.80		0.0002	

Control only  
M98Z

Model	n	R	St. err.	Difference	95% CI	p-value	AIC
Codominant	153 109	19.44 19.8	0.12 0.16	0.00 0.36		0.0019 0.02314	1088.9
Dominant	153 136	19.44 19.95	0.12 0.14	0.00 0.51	( 0.14 - 0.88 )	0.0068 0.03934	1092.1
Recessive	262 27	19.59 20.57	0.1 0.29	0.00 0.99	( 0.36 - 1.61 )	0.0023 0.02329	1090.1
Overdominant	180 109	19.61 19.8	0.12 0.16	0.00 0.19		n.s.	
Log additive	---	---	---	0.49	( 0.21 - 0.76 )	0.0006 0.0162	1087.8