

ADAPTIVE SIGNIFICANCE OF PELVIC GIRDLE LOSS IN THREE SPINE
STICKLEBACK

by

Mirjam Barrueto

B.Sc., University of Zürich, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Zoology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2009

© Mirjam Barrueto, 2009

Abstract

Pelvic girdle loss has evolved repeatedly in freshwater stickleback. In many cases it is caused by a mutation in a regulatory sequence that controls expression of the major pelvic girdle locus (*Pitx1*) in the prospective pelvic region. Pelvic girdle loss can spread rapidly through populations, but its adaptive significance is not well understood. I experimentally tested potential ecological mechanisms for pelvic girdle loss in juvenile threespine stickleback. I carried out two insect predation experiments to measure selection on pelvic spine length and body size, and to measure the effect of the pelvic girdle on survival. In the first experiment, the pelvic spines of all fish were clipped to varying degrees to artificially create variation in spine length. For the second experiment, I created hybrid backcross families, where 50% of the offspring expressed pelvic girdle loss. In both experiments, the fish were exposed to predation by a common aquatic insect predator. Selection tended to favour shorter spines, increased body size and loss of the pelvic girdle. A third experiment measured growth rates in juvenile fish with and without a pelvic girdle, to test for a potential growth rate advantage of the *without* phenotype. On average, *without*s exhibited a nonsignificant trend for higher growth rates compared to *with*s.

As the power of the predation experiments was low, I performed meta-analyses to combine my data with previously published experiments. Across these experiments, insect predators selected for shorter pelvic spine length as well as increased body size. There was no mean positive or negative effect of the pelvic girdle on survival in the face of insect predation, but the confidence interval for this result was wide, and further studies are required.

My findings suggest that predation may drive pelvic girdle loss in juvenile stickleback by means of selection against correlated traits, such as long pelvic spines, rather than selection against the pelvic girdle itself. I did not detect any significant association between pelvic girdle loss and increased growth rates. It thus remains undetermined if selection for increased body size acts as an additional selection factor for pelvic girdle loss.

Table of Contents

Abstract	ii
Table of Contents	iii
List of Tables	v
List of Figures	vi
Acknowledgements	vii
1 Introduction	1
1.1 Morphological divergence in adaptive radiation	1
1.2 Genes of large effect in phenotypic evolution	2
1.3 Pelvic reduction in threespine stickleback	2
1.4 Mechanisms of pelvic reduction in stickleback	4
1. Predation	4
2. Differences in growth rates	6
3. Calcium limitation	7
2 Methods	10
2.1 Choice of study populations and experimental crosses	10
2.1.1 Insect predation experiment 1	10
2.1.2 Insect predation experiment 2	10
2.1.3 Growth rates experiment	11
2.2 Experimental procedures and analysis	12
2.2.1 Insect predation experiment 1	12
2.2.2 Insect predation experiment 2	14
2.2.3 Growth rates experiment	16
3 Results	19
3.1 Insect predation experiment 1	19
3.2 Insect predation experiment 2	19
3.3 Growth rates experiment	20
4 Discussion	21
4.1 Insect predation hypothesis	21
4.2 Growth rates hypothesis	22
4.3 Comparison with other data	23
4.3.1 Selection on pelvic spine length	24

4.3.2	Selection on standard length	24
4.3.3	Selection on presence or absence of the pelvic girdle	25
4.4	Conclusions	27
References	45
A	Appendix A	46
A.1	Mean pelvic spine length in predation experiment 1	46
A.2	Mean standard length in predation experiment 1	47
A.3	Mean pelvic spine length (residual) in predation experiment 1	48
A.4	Growth rates of the two pelvic girdle phenotypes	49
B	Appendix B	50
B.1	Methods for comparisons among experiments	50
B.2	Mean selection differentials	50
B.3	Mean log odds ratios	51
B.4	Power of pelvic girdle predation experiments	53
C	Appendix C	57
C.1	Animal care certificate	57

List of Tables

1	Standardized selection differentials and significance tests (insect predation experiment 1)	29
2	Comparisons among insect predation experiments	30
3	Aquatic insect predation experiments	31
A.1	Mean pelvic spine length in predation experiment 1	46
A.2	Mean standard length in predation experiment 1	47
A.3	Mean pelvic spine length (residuals) in predation experiment 1	48
A.4	Growth rates of the two pelvic girdle phenotypes	49
B.1	Results of the experiments from Reist (1979)	55
B.2	Results of the experiments from Reimchen (1980)	56

List of Figures

1	Insect predation experiment 1	32
2	Insect predation experiment 2	33
3	Growth rates of all families, including 100% of all fish	34
4	Growth rates of all families, only including the largest 50% of all fish	35
5	Mean standardized selection differentials for pelvic spine length	36
6	Mean standardized selection differentials for standard length	37
7	Mean log odds ratios from the different insect predation experiments	38
B.1	Power of insect predation experiments	54
C.1	Animal care certificate	58

Acknowledgements

Many people helped me with my thesis. I'm especially grateful to Sahriar Kabir, who was a great help and good company during a long summer with even longer hours while carrying out many experiments at once. Kerry Marchinko, Rowan Barrett, Joey Courchesne, Pat Tamkee, Sean Rogers, Anne Dalziel and Matt Arnegard of the Schluter lab always made time to answer all sorts of urgent questions. Thanks also to Jason Clarke, Antoine Paccard and Gina Conte for lots of moral support, in addition to all the volunteers, assistants and summer students that helped keep the fish alive and happy. I would further like to thank the many people in the Zoology and Botany departments who taught me about evolution and (ecological) speciation. A special thanks to my committee, Chris Harley and Jon Shurin, who gave advice and made the committee meetings and writing of the thesis very pleasant. Above all, I'm thankful to Dolph Schluter for being a great and patient supervisor and for giving me the opportunity to come here! I very much appreciated the freedom I had to make mistakes, learn from them and to find my own questions and answers, while still being guided and supported all along. The importance of proper experimental design and the necessity of knowing R will probably stick in my mind for a long time!

1. Introduction

Divergent natural selection, acting on traits of populations inhabiting different environments or on traits within a population, can lead to phenotypic and ecological differentiation (Schluter, 2000). According to the concept of adaptive radiation, periods of such divergence may exhibit increased rates of speciation and are thus a major focus of biodiversity research. Many cases of ecological selection have been studied and in some cases the underlying ecological mechanisms have been found, but studies that quantify selection on naturally occurring genetic variation are rare. Additionally, research on natural selection leading to genetic and phenotypic diversification of organisms has mainly focused on the role of competition for resources and mates (Vamosi, 2005; Langerhans, 2006). Despite having an acknowledged importance in population and community ecology, the role of predation on the evolution of diversification has only recently received more theoretical and empirical attention (e.g. Reznick and Endler, 1982; Abrams, 2000; Rundle et al., 2003; Nosil and Crespi, 2006; Meyer and Kassen, 2007). This thesis attempts to estimate the adaptive significance of a naturally occurring change in a morphological trait, which is strongly determined by a single locus of large effect and might be mediated by predation pressure.

1.1. Morphological divergence in adaptive radiation

Divergent natural selection among populations of the same species, inhabiting differing environments, can be depicted by fitness landscapes, where each environment might feature distinct adaptive peaks and valleys. A consequence of the idea of divergent natural selection is that intermediate phenotypes are situated in the fitness valleys and populations with intermediate mean phenotypes are pulled towards regions of higher fitness around them. The ecological theory therefore holds that divergent natural selection is responsible for the correlation between phenotype and environment in an adaptive radiation (Schluter, 2000). In most known examples, natural selection acts on morphological traits. Other traits such as physiological, life history or behavioural characteristics have been studied less intensely, presumably because morphological traits are often easier to identify and measure (Kingsolver et al., 2001). A good example of a morphological trait un-

der natural selection is bill shape in red crossbills (*Loxia curvirostra* complex) (Benkman, 2003). In these species, bill morphology is directly linked to the birds' ecology. Bill depth has a high heritability and is a primary target of selection. It is associated with handling efficiency of food items, and thus determines the birds' use of food resources and ultimately survival in different environments. Based on this knowledge, individual birds' fitness can be predicted from their bill phenotype (Benkman, 2003). Most examples from nature, however, still lack such detailed understanding of mechanisms and fitness consequences.

1.2. Genes of large effect in phenotypic evolution

Until very recently, most mathematical theories predicted the number of genes contributing to phenotypic evolution to be large, with each locus having an infinitesimal effect on the phenotype (Orr, 2005). This assumption forms the basis of quantitative population genetics. Newer empirical data, however, tells a different story. A recent review of the genetics of species differences found the number of quantitative trait loci (QTL) involved in the differences to be highly variable among the studies, ranging from cases with very few QTL of large effect to cases with many loci that each contribute only very little (Orr, 2001).

1.3. Pelvic reduction in threespine stickleback

Pelvic girdle loss in freshwater threespine stickleback (*Gasterosteus aculeatus* species complex) is one of the best-known examples of morphological adaptation conferred by one gene of very large effect, complemented by a handful of small modifier loci (Shapiro et al., 2004). In the ancestral marine stickleback, the pelvic girdle and the pelvic spines form the pelvic skeletal structures that protect the fish from gape-limited, soft-mouthed predators. The girdle consists of two ventral plates that extend up the sides of the fish. The two pelvic spines articulate with the underlying pelvic girdle, and can be locked erect to increase the effective diameter of the fish, as well as cause damage to soft mouthparts of predators (Hoogland et al., 1957; Reimchen, 1983).

In many cases, when the marine threespine stickleback, inhabiting coastal waters all around the

Northern hemisphere, colonized freshwater habitats, they underwent reduction and in some cases loss of the pelvic structures. There are two forms of pelvic reduction in stickleback: *Size reduction*, where most or all parts are still present but size-reduced to varying extents, and *complete loss* (sometimes called *extreme reduction*), where the girdle and spines are either lost altogether, or the spines are lost and the girdle plates reduced to tiny bony vestiges (Bell, 1987). The genetic basis of pelvic reduction is comparatively well described. It has been surprisingly difficult, however, to identify the ecological mechanisms that underlie this dramatic morphological adaptation. This question is the focus of my thesis.

Fossils of threespine stickleback from the late Miocene document that pelvic girdle loss can happen rapidly; in one case the pelvic girdle was lost from a population within a few thousand generations (Bell et al., 2006; Hunt et al., 2008). Most extant stickleback populations that show pelvic reduction are young, occur in recently deglaciated regions and are restricted to lacustrine environments. The different populations have often evolved separately from local marine populations (Hagen and McPhail, 1970). When populations with pelvic reduction are ordered by the amount of reduction, they typically have a continuous distribution of phenotypes, often including fish with complete pelvic structures. The frequency distributions of the different phenotypes, however, are highly variable among populations (Bell, 1987). Populations that exhibit complete loss of the pelvic structures are rare, but can be found across the whole natural range of *G. aculeatus*: in Scotland (Coyle et al., 2007), Iceland (Shapiro et al., 2004), Alaska (Cresko et al., 2004), and Paxton and Little Quarry Lakes in British Columbia, Canada (McPhail, 1992). The two Canadian populations of fish without a pelvic girdle (“*benthics*”) are special insofar as they both coexist with a second ecological species of threespine stickleback (“*limnetics*”), which have retained complete, but size-reduced, pelvic structures (McPhail, 1992).

Pitx1, the stickleback orthologue of a human homeodomain transcription factor involved in hind limb development, mediates pelvic girdle loss in the Paxton Lake benthic species. Complementation tests also indicated that *Pitx1* is responsible for pelvic girdle loss in the Icelandic population. The coding region of the Paxton benthic *Pitx1* gene itself is unchanged from the ancestral marine sequence. Instead, regulatory changes cause the absence of *Pitx1* expression in the prospective

pelvic regions (Shapiro et al., 2004). Similar genetic mechanisms, involving *Pitx1* and in some cases minor modifier loci, were found in the other known populations of stickleback with extreme pelvic reduction (Cresko et al., 2004; Shapiro et al., 2004; Coyle et al., 2007). In contrast, in one population that only shows size reduction but no complete loss of the pelvic structures, *Pitx1* is not involved in reducing pelvic girdle or spine length (Peichel et al. (2001)).

1.4. Mechanisms of pelvic reduction in stickleback

Despite the advances in understanding the genetics of pelvic reduction in threespine stickleback, knowledge of the ecological mechanisms that lead to and maintain the within- and between-population variation of this phenotype, is still patchy. The fact that pelvic reduction has occurred repeatedly and in some cases by the same molecular mechanisms is strong evidence for parallel natural selection being the cause. In addition, the differences in armour traits between the limnetic and benthic species pairs in Paxton and Little Quarry Lakes are persistent over time, despite the potential for interbreeding (McPhail, 1992). This suggests that selection is maintaining the pelvic phenotypes of the two species pairs.

There are currently three hypotheses of ecological mechanisms leading to pelvic reduction in three-spine stickleback. The adaptive significance of pelvic girdle loss specifically, however, is not well understood yet.

1. Predation

The amount of external body armour of stickleback is correlated with the type and abundance of predators (Vamosi, 2003; Marchinko, 2009). In marine and some freshwater habitats, stickleback are intensely preyed upon by gape limited piscivorous fish and birds (Reimchen, 1994), which are only able to ingest prey they can swallow whole. These sticklebacks are heavily armoured with robust pelvic and dorsal spines. Compared to the ancestral marine habitat, gape limited predators are less abundant in many freshwater habitats, and, especially in lakes and ponds, may even be absent (Reimchen, 1994). Instead, aquatic invertebrates predate upon juvenile and small adult sticklebacks. In contrast to piscivorous fish and birds,

insect predators do not swallow their prey whole, but catch the prey by grabbing them. They then either inject a toxin to paralyze and liquefy the fish (e.g. *Notonecta sp.*) or bite off pieces from the captured but live prey (*Aeshna spp.* nymphs, *Dysticus spp.*). Up to a size of about 15 mm the fish are also vulnerable to cannibalism by adult stickleback. This has been shown to drive young fry into seeking refuge in vegetation, thus increasing the risk of predation by insects (Foster et al., 1988). Under these predation regimes, stickleback often have short or no spines and a reduced or absent pelvic girdle (Reimchen, 1980; Bell et al., 1993). One often-cited but not explicitly tested hypothesis by Reimchen (1980) states that the pelvic spines and pelvic girdle might provide structures for invertebrate predators to grab and hold on to, and are thus selected against. Direct observations of predatory attacks are needed to determine the role of the different pelvic structures during the attack and capture phases, and to estimate their effect on escape probabilities (Vermeij, 1982; Langerhans, 2006).

Relaxed selection for pelvic armour due to absence of gape-limited predators, combined with selection against heavily armoured fish by aquatic invertebrate predators is thus a particularly plausible mechanism for pelvic reduction. Indirect evidence for a role of aquatic invertebrates in spine number reduction in juvenile stickleback is provided by Reimchen and Nosil (2002), who found a correlation between season and spine number. Vamosi (2002) conducted predation experiments in littoral arenas that mimicked the habitat of benthics, where groups of juvenile Paxton Lake benthics and limnetics were subjected to predation by backswimmers (*Notonecta sp.*), dragonfly larvae (*Aeshna spp.*) and adult benthic stickleback. In this experiment, juvenile limnetics, which have significantly more pelvic armour than benthics, had lower survival in the presence of backswimmers and adult stickleback than benthic juveniles, but not in presence of dragonfly larvae. However, since the two species differ widely in many traits, it was not possible to determine whether the observed differences in survival between benthics and limnetics were caused by morphological, behavioural or other traits. In another study by Marchinko (2009), juvenile F₂ hybrid fish from a cross between marine (large robust pelvic girdle) and freshwater threespine stickleback

from two populations (one pelvic-reduced and the other lacking a pelvic girdle) were subjected to predation by dragonfly larvae and backswimmers in pond enclosures. Those fish with the largest body size, shortest dorsal spines and shortest pelvic girdle length tended to survive best. Surprisingly, individuals completely lacking a pelvic girdle exhibited increased mortality.

Both these studies confirmed a role for invertebrate predation in armour reduction. They were, however, not designed to eliminate covariance of several armour traits, making it impossible to determine which traits are targeted by selection. They also did not control for potential negative epistasis arising from crossing two species, which are, although closely related, highly morphologically divergent (Marchinko, 2009). Alleles derived from one of the species might negatively affect the fitness of the hybrid individual, if expression of these alleles in the new genetic background leads to detrimental genetic interactions. Insect predation experiments on stickleback have never focused on the individual morphological traits such as the pelvic spine or girdle. The role of individual structures in the evolution of pelvic reduction has yet to be determined.

2. Differences in growth rates

In stickleback, individuals with fewer lateral armour plates grow faster than fish with more lateral plates (Marchinko and Schluter, 2007; Barrett et al., 2008). Lateral plates are, like the pelvic girdle, external bone structures. It is plausible that the presence or absence of a pelvic girdle has an influence on a fish's growth rate in a similar way. Faster growth is often thought to be generally advantageous for individual organisms. In freshwater stickleback, advantages of faster growth in juvenile fish may include a more rapid escape from those size classes most vulnerable to predation by aquatic invertebrates and adult stickleback (e.g., Marchinko, 2009; Foster et al., 1988), enhanced over-winter survival (Curry et al., 2005) and earlier breeding (Schultz et al., 1991). Yet there are also potential disadvantages to fast

growth in juvenile fish, including increased developmental instability (Arendt, 1997; Robinson and Wardrop, 2002), which has repeatedly been associated with fitness loss (for a review: Møller (1997); but see: Clarke (1998)). Rapid growth is usually at least partly genetically determined, and might decrease an individual's ability to tolerate nutrient stress (Arendt, 1997). It can also lead to fluctuating asymmetry in phenotypic structures, which may have implications on predator-prey interactions and subsequent niche segregation patterns of the prey, complicating the picture further (Reimchen and Bergstrom, 2009).

If pelvic reduction indeed allows faster growth rates of juvenile fish, it would not necessarily be under positive selection. In environments where gape-limited predators select for robust pelvic armour, benefits of faster growth through pelvic reduction would likely be outweighed by increased mortality of those fish with reduced or absent pelvic structures. In habitats with reduced predation by gape-limited predators, however, the advantages of increased growth rates might lead to selection for pelvic reduction or pelvic girdle loss. This hypothesis has never been tested experimentally.

3. Calcium limitation

Giles (1983) proposed that calcium limitation might pose strong selection pressure against armour development (*calcium limitation hypothesis*) in calcium poor habitats, such as many freshwater lakes, since teleost fish absorb much of the calcium necessary for skeletal development from ambient water (Simkiss, 1974; Koenings et al., 1987; Peterson and Martin-Robichaud, 1986). Bell et al. (1993) showed that in 179 lakes sampled around Cook Inlet, Alaska, low calcium concentrations are associated with pelvic reduction, but only when predatory fish are absent. It was also found that pelvic reduction usually occurs at calcium-concentrations below a certain threshold (12mg/l). Pelvic reduction was not inevitable, however, even at low calcium concentrations and in the absence of predatory fish. Additionally, there are examples of severely pelvic reduced stickleback populations that evolved in lakes with high calcium concentrations (e.g. Paxton Lake, which has a marl bottom (Larson, 1976)). The authors concluded that calcium availability and predatory fish cannot fully pre-

dict pelvic reduction, and suggested that the effects of other abiotic factors (low iron and phosphate availability and the low ionic strength of freshwater) as well as biotic factors (insect predation) be investigated (Bell et al., 1993). The calcium hypothesis has not been tested experimentally.

My thesis tested combinations of these three proposed mechanisms. I estimated selection on pelvic reduction in benthic fish from Paxton Lake with a predation experiment on two pelvic phenotypes, and measured growth differences between pelvic phenotypes at low ambient calcium concentrations. I also estimated the strength of selection on pelvic spine length with a controlled predation experiment that manipulated length of pelvic spines, thus minimizing the influence of confounding variance. Previous predation experiments on stickleback have mostly been carried out on the offspring of crosses between highly diverged species, with the resulting hybrids expressing high variation in the traits of interest, which selection was then allowed to act upon. The advantage of this approach is that the effects of many different morphological features can be tested within one experiment. In addition, if several traits with analogous functions (e.g. armour traits) are linked, total selection can be detected more readily. The same approach, however, might not allow the contributions of all individual morphological features to be singled out. Multiple regression techniques can often identify the traits that are under strongest selection, assuming that all important characters were measured (Lande and Arnold, 1983). If this assumption is not met and the crosses exhibit variation in other, unmeasured traits, the outcome of the experiment is confounded and traits of small effect are impossible to detect.

I conducted the two selection experiments under semi-natural conditions. In the first insect predation experiment (experiment 1), I manually shortened the pelvic spines of fish that naturally possess a full pelvic girdle with robust, long spines. As a test of the predation hypothesis, I then assessed if individuals with longer pelvic spines were more vulnerable to predation by one major invertebrate predator. For this, families of manipulated fish were exposed to backswimmers (*Notonecta* sp.), and the spine lengths of the surviving fish were compared to fish in control treatments without predators. In the second insect predation experiment (experiment 2) I also measured survival in

the face of invertebrate (backswimmer) predation, but using backcrossed hybrid families, in which half of the fish had complete pelvic girdles and the other half had no pelvic structures. Here I compared survival of fish with a pelvic girdle with those fish lacking one (predation hypothesis). In addition, in both insect predation experiments, I measured selection on body size to test for a potential benefit of larger body size that might arise from faster growth.

Finally, as a direct test of the growth rates hypothesis, I carried out a third experiment that examined growth rates under very low calcium levels in the ambient water. I measured within-family differences in juvenile growth rates in families where half the fish had a pelvic girdle and half the fish had none.

2. Methods

2.1. Choice of study populations and experimental crosses

2.1.1. Insect predation experiment 1

In this experiment I used stream fish collected from Salmon River, BC, at two locations (49° 9'13"N; 122°35'25"W and 49° 9'35"N; 122°35'22"W), which displayed a robust pelvic girdle and large pelvic spines. They also were of the low lateral plate morph. For a direct test of the effect of the pelvic spines on survival in the face of backswimmer predation, I shortened the pelvic spines of all fish to various degrees to test for selection on spine length.

2.1.2. Insect predation experiment 2

To test whether presence or absence of the whole pelvic girdle (including the spines) had an impact on survival in the presence of insect predation, I created backcrosses between pure Paxton benthics and F₁ benthic – limnetic hybrids. Some of the benthic fish for this experiment were caught with minnow traps at Paxton Lake itself (49°36'43"N, 124°01'57"W), on Texada Island, BC. Others were sampled from an experimental pond on campus at the University of British Columbia, BC (49°15'9"N, 123°14'13"W) which had been stocked with wild caught Paxton benthics in 2005 and was then left to reproduce naturally. The F₁ benthic-limnetic hybrids came from two crosses that were made in 2007 from wild caught Paxton benthic and limnetic parents, and were raised in the lab. The F₁s were presumed to be heterozygous at the *Pitx1* locus, carrying one pelvic-present allele (from the limnetic parent) and one pelvic-absent allele (from the benthic parent). The pelvic-present allele is dominant (Shapiro et al., 2004) and as expected, all F₁ hybrids had a full pelvic girdle. The pure benthics used in making the backcrosses had no pelvic structures at all. In the resulting backcrosses, about half of the offspring thus had a full pelvic girdle and about half lacked a pelvic girdle. By using these backcross families, the genetic background of the fish was three quarters benthic and the influence of other genes that were not tightly linked to *Pitx1* was mini-

mized.

2.1.3. Growth rates experiment

The growth experiment was conducted on the offspring of crosses between two pelvic girdle morphs of Paxton benthics. The large majority of Paxton benthics have no pelvic structures. Up to 20 percent of fish in Paxton Lake that otherwise appear to be benthics, have pelvic structures (McPhail (1992); M. Barrueto, unpublished data), ranging from the presence of only a single plate to presence of a full pelvic girdle with spines. I collected such benthics (from here on referred to as "benthics_{PG}") for the growth rates experiment. Adult benthics and limnetics are easily distinguished visually and I used this method to classify benthic-like individuals with a girdle as benthics_{PG}. It is possible that benthics_{PG} have received their pelvic girdle through introgression of limnetic alleles into the benthic population and might be first or advanced generation hybrids, but this is not confirmed. For the experiment I chose fish with a full benthic girdle and a benthic-like appearance to increase the probability that they were later-generation backcrosses with a minimum of other limnetic genetic material. All benthics_{PG} that were used for the crosses were presumed to be heterozygous for the limnetic *Pitx1* allele. This presumption was corroborated by the pelvic girdle distribution of their offspring: as expected, half the fish had a full pelvic girdle and half had no pelvic structures. I crossed the wild benthics_{PG} from Paxton Lake with benthics from either the wild (Paxton Lake) or the pond (see methods of insect predation experiment 2). Because of a shortage of benthics_{PG} I used several of them to make more than one cross (Table A.4), but with different partners for each cross. From the 13 crosses that were made for this experiment, nine were analyzed (see below).

2.2. Experimental procedures and analysis

2.2.1. Insect predation experiment 1

Fertilization was done artificially, following standard methods (e.g. Marchinko, 2009, Supporting Information). The eggs of a highly gravid female were obtained by gently squeezing them from her abdomen. The male was euthanized in MS222 and his testes removed. Each male only fertilized one female. I thoroughly mashed both testes in a few drops of water to release the sperm, and added the mix to the eggs. Most crosses were done at the trapping sites and the fertilized eggs transported to the lab in water filled, cooled and oxygenized Falcon™ tubes. In the lab, the eggs were placed in hatching cups and suspended above an air bubbler in 100 litre tanks of dechlorinated tap water. In order to provide the fish with ample calcium for optimal growth, coarse washed limestone was added to the tanks, as well as a tablespoon of powdered dolomite. I added dissolved methylene blue to the water (enough to lightly colour the water) to prevent fungal and bacterial infections of the eggs. After approximately eight days the embryos hatched. They were fed live microworms (*Panagrellus* sp.) and newly hatched brine shrimp twice daily. When the fish reached a size of 14 to 15 mm, I randomly split each family into two groups and shortened the pelvic spines of one group with nail scissors, as close to the base as possible. I clipped only the tips of the spines in the remaining fish. The fish were then kept in separate tanks overnight. I added tea tree oil (Melafix® by Aquarium Pharmaceuticals) to the water to aid recovery from the injury.

The following day I randomly split the surviving fish of each treatment into two groups. I then randomly combined the clipped and tip-clipped subgroups within each family into one predation treatment group (predators added) and one control group (no predators added), and brought them to the experimental area. The experiments took place in 20 189-litre plastic tanks (Rubbermaid® *Jumbo Roughtote*). Each tank had a sandy bottom (sand from nearby ponds), and was filled with pond water. I added *Chara*, dead branches and rocks in order to create a realistic habitat, and covered the tanks with mosquito mesh, to prevent other insects laying eggs in the tanks. There were no predators in the tanks prior to the start of the experiment, but small pond organisms were present in the water. I added the predation treatment and control groups to one tank each. Every

trial consisted of fish from one single family only.

I assigned ten ‘blocks’ comprising two neighbouring tanks, one treatment (insects added) and one control (no insects added), and generated a random order for block occupation of the trials. Assignment of experiment and control tanks within a block was randomized for each trial. The number of fish per tank ranged from 20 to 41 (predation treatment) and 19 to 40 (control treatment). I left the fish in the tank overnight to acclimatize and fed them one 3.5oz cube of frozen *Daphnia*. 24 hours later I removed and stored any dead fish, which were easy to spot, and added backswimmers that had been caught the same day in a nearby pond, to the experimental tanks. Following Marchinko (2009), I added 0.6 backswimmers for each fish. Every three days, I counted the surviving fish in the predation treatment tank, and stocked the tank with new backswimmers, if any were missing. I fed the fish one cube of frozen *Daphnia* every three days. When around 50 percent of the fish in one predation treatment tanks had died, I collected all remaining fish of this trial, as well as the control fish, and euthanized them in MS222 for storage in 95% EtOH. The very first trial, in contrast to all the other trials, only showed mortality of fish that had most of their spines removed (compared to those that only had the tips clipped). I had practiced clipping the pelvic spines without killing the fish on many families before starting the experiment. Mortality due to clipping in “tip-clipped” fish was very low from the start, but in the beginning, a large percentage of fish that had their “whole” spines clipped, died within 24h of the clipping. Thus it was suspicious when, in the very first trial, only fish with their whole spines clipped exhibited mortality, and thus the trial was excluded from subsequent analysis.

Insect predation experiments on juvenile stickleback have previously been found to result in selection for larger body size (Marchinko, 2009). Hence, body size had to be taken into account in the analysis of spine length, since pelvic spines grow with body size. Because I had manipulated spine length, however, the correlation of spine length and body size was not expected to be as strong as in un-manipulated fish. Standard length is highly correlated with body size in fish (Baumgartner et al., 1988), and thus was used as a proxy for body size. Standard length was determined as the Euclidian distance between the tip of the snout and the beginning of the caudal fin, measured on digital photographs of each individual fish using the program tpsDig2 (Rohlf, 2006). Pelvic spines

were measured from the distal tip to the point of articulation with the pelvis. I used callipers and 25x magnification and measured to the next 0.01 mm. The lengths of left and right spines were summed.

I used repeatability to assess reliability of the measuring method for pelvic spine length. Repeatability (r) compares the proportion of variance in a trait that occurs among individuals rather than within individuals when multiple measures on the same individual are taken (Lessels and Boag, 1987). The variances were obtained by fitting a linear mixed-effects model to the spine length data of one randomly chosen trial (trial 16, 31 individuals), where every individual had been measured twice, on different days. Repeatability (r) for the spine length measure was then calculated following Lessels and Boag (1987), and was 0.991 (i.e. 0.9% of the variation found in spine length was due to measurement error). The standard deviation for the repeated measurements was 0.140 mm. I estimated selection on total pelvic spine length, standard length, and size-corrected pelvic spine lengths. Standardized selection differentials (i) were obtained separately for each trial as in Marchinko (2009), using equation (6.1) in Endler (1986),

$$i = \frac{\bar{X}_a - \bar{X}_b}{\sqrt{\text{var}_b}} \quad (1)$$

with \bar{X}_a and \bar{X}_b representing the mean trait values of fish in the predation and control treatment, respectively, and with var_b as the trait variance of the control treatment. The selection differentials of the size-corrected spine lengths were obtained by using the residuals from an ordinary least squares regression of total pelvic spine length on standard length, combining all surviving fish from all trials.

2.2.2. Insect predation experiment 2

The backcross families were raised in the same way as the fish in insect predation experiment 1. Once the families had reached a mean size of 14 to 15 mm, I randomly split each family into two subgroups (predation and control treatment). Only fish smaller than ~16 mm were included

in the trials, since larger fish are usually not preyed upon by backswimmers (Foster et al., 1988). In one trial I combined two families due to small family size; I split the families separately and then combined the subgroups. The experimental procedure was essentially the same as in the first predation experiment. The number of fish in each experimental tank ranged from 25 to 39 (predation treatment) and 24 to 37 (control treatment). The experiment itself consisted of adding 0.6 backswimmers per fish to the predation treatment, and measuring survival of the different phenotypes. Every three days, I counted the surviving fish in the predation treatment tank, collected the dead fish and, if necessary, stocked the tank with new backswimmers. When about half of the fish in the predation treatment had died, I collected all the remaining fish, as well as the control fish, and euthanized them in MS222 for storage in 95% EtOH. In order to keep experimental conditions as uniform as possible, I excluded all trials from analysis that showed less than 30% mortality in the predation treatment and lasted for longer than 14 days. Also, if there was more than 10% mortality in the control treatment, I excluded the whole trial from analysis.

The dead fish were fixed in formalin for three days. After rinsing, they were stained for 24h in an Alzarin Red to make the girdle, plates and spines more visible. When only the ovoid vestiges (see Bell, 1987) were present, I considered the fish to be of the ‘no pelvic girdle’ phenotype. Those fish that had an asymmetric pelvic girdle (structures not complete on both body sides) were included in the ‘pelvic girdle’ category. I photographed the stained fish with a digital camera and measured their standard length in millimetres using tpsDig2 landmarks, as in predation experiment 1. The starting frequencies of fish with and without pelvic girdle in the trials were not known. However, because the assignment of fish to predation treatment and control groups had been random, it could be assumed that at the beginning of the experiment the two groups displayed the same proportion of fish with and without a pelvic girdle.

To determine survival of the two categories, I calculated the proportion of surviving fish without a pelvic girdle in the predation treatment and compared it to the same measurement of the control group. I tested for differences in survival between the two pelvic girdle categories with a paired *t*-test and calculated the 95% confidence interval to obtain an estimate of effect size of pelvic girdle on survival. As a second measure, I calculated the log odds ratio for survival of fish without a

pelvic girdle, following the procedures of Whitlock and Schluter (2009).

Selection on body size was quantified by calculating standardized selection differentials for standard length, following the same methods as in insect predation experiment 1.

2.2.3. Growth rates experiment

I raised the fish in dechlorinated tap water, but without added limestone or other calcium-containing water-hardener, to measure growth under calcium limitation. Calcium concentrations of tap water in the City of Vancouver averages between 0.95 and 1.81 mg/l, depending on the source of the water (*Drinking water quality. 2008 Annual report. City of Vancouver*; available on <http://www.vancouver.ca>), which is much lower than the “calcium threshold” of 12mg/l below which Bell and colleagues (1993) observed pelvic reduction. The amount of calcium that was obtained through the food was not monitored. Calcium was not an added ingredient in any of the food used during the experiment (for ingredient lists, see <http://www.hikariusa.com/products/bio-pure/> for *Daphnia*, bloodworms and mysis shrimp). *Daphnia* spp. have calcium contents ranging from 2.8 to 7.7% of dry weight (see Alstad et al., 1999) and were the only food that was administered to the fish during the first seven to ten weeks of the measurement period (see below). Calcium availability was, therefore, likely to be rather limited.

After hatching, the fish were raised for approximately five weeks in one 100-liter tank per family. Densities ranged from around 40 to 120 fish per tank during that time. All the tanks were situated in an environment chamber with a constant air temperature of 20°C. The tap water temperature was 16°C. For the whole duration of the experiment, the chamber was set to a 14 hours light, 10 hours darkness cycle. The fish were fed live brine shrimps twice daily and microworms once daily. When they reached a mean standard length of 12 to 14 mm, with the pelvic girdle and pelvic spines clearly visible, I divided the families: 30 fish were kept in the experimental tanks, 15 individuals with (*withs*) and 15 individuals without (*without*s) pelvic girdle. The remainder of the family was kept in a second backup tank.

Every 21 days (+/- 2 days) I measured standard length of each individual, holding the fish in the hand and using a small ruler. This method was developed by Marchinko (2009) and was very efficient to process lots of fish, but not extremely precise, as the size could only be measured to the nearest half-millimetre. Due to their small size the fish could not be marked individually, so the measurements were averaged, for each family, at each measurement, separately for *withs* and *without*s. The fish were first measured at the age of five to eight weeks, and measurement continued until the mean size per family reached or exceeded 30 mm. This was generally 25 to 28 weeks post-hatching. Fish that died within the first six weeks of the experiment were replaced with haphazardly selected fish of the same pelvic phenotype from the backup tanks, which had received the same feeding and water change treatments as the experimental tanks. After six weeks, the sizes of fish in some replacement tanks and experimental tanks began to diverge, with the replacement fish growing faster, and subsequently, dead experimental fish were not replaced anymore. This divergence was probably due to lower densities in some of the replacement tanks, since otherwise all tanks were treated in the same way. In order to maintain good water quality, a third of the water of each aquarium was changed every three days using tap water, Amquel[®] was added at regular intervals to detoxify the water and the pH was maintained between 6.8 and 7.4 by adding sodium bicarbonate. Using food cubes allowed standardization of food administration. Each *Daphnia* cube weighed 3.5oz. Each cube of *worms* consisted of 10 g of a 2:1 mixture of chopped frozen bloodworms and mysis shrimps (all food was of the brand Hikari[®] Bio Pure[®]). Every tank received three cubes daily, once the measurements had started. Until week 14 I fed each family frozen *Daphnia* cubes. From week 15 on the juvenile fish were fed a mix of frozen *Daphnia*, bloodworms and mysis shrimp. The proportion of worm cubes in the diet was increased as the fish grew (one cube of worms every second day during weeks 15 to 22; two cubes of worms every second day during weeks 23 to 26; two cubes of worms daily during weeks 27 and 28). The fish were fed to satiation once daily, and two hours after feeding the food remains were removed from the tanks.

Because I had replaced dead fish in the beginning of the experiment, and because there was some mortality in most families throughout the whole experiment, I could not calculate overall growth

rates by simply using starting (week five) and final (around week 28) body size. Instead, I calculated mean growth rates [mm/day], separately for the two pelvic girdle categories of each family, from the difference in mean standard lengths of two consecutive measurements. Each family was measured up to eight times in 21 day intervals. From these growth rate measures, I calculated an overall mean growth rate for each phenotype and family. I stopped measuring families and excluded them from analysis if the number of fish dropped below 20 individuals early in the experiment (before the fifth measurement), to avoid the introduction of density as a confounding variable. Given these restrictions, a total of nine families (from an initial 13 families) could be analyzed. An estimate of the effect of the pelvic girdle on growth rates was obtained through the 95% confidence interval for the difference in growth rates. I formally tested for such a difference by performing a paired *t*-test.

3. Results

3.1. Insect predation experiment 1

Under aquatic insect predation, juvenile stickleback that had shortened pelvic spines tended to survive best (Table 1, Figure 1). This trend was close to statistically significant for the standardized selection differentials of the residuals (pelvic spine length – standard length regression), which corrected spine length for body size (mean= -0.163; standard error= 0.073; $t = -2.252$; $df = 6$; $P = 0.065$; Table 1; *Pelvic spine length (residuals)*). The trend was still present, albeit less pronounced, when spine length was not corrected for body size (Table 1; *Pelvic spine length*). Standard length by itself was under no detectable directional selection during the experiment (Table 1).

3.2. Insect predation experiment 2

In the presence of insect predation, fish without a pelvic girdle tended to have a higher probability of survival than fish with a pelvic girdle. Among the surviving fish, the mean proportion of individuals without a pelvic girdle was 10.9% higher in the predation treatments compared to the control treatments (Figure 2). This result was nonsignificant, however ($t = 1.012$, $df = 8$, $p = 0.341$).

Similarly, the mean log odds ratio of survival for fish without a pelvic girdle was positive ($=0.587$, indicating an 80% increase in survival probability for fish without a pelvic girdle compared to fish with), with a standard error of 0.542 and a 95% confidence interval from -0.475 to 1.650.

Surviving fish in the predation treatment tended to be longer than those in the control treatment (mean difference of 3.82%), though the trend was not statistically significant ($t = 1.992$, $df = 8$, $P = 0.082$). The mean selection differential for standard length was 0.337 with a standard error of 0.201 (full results in Appendix A). There was no significant size difference between the pelvic phenotypes among the surviving fish in the control treatments (fish without a pelvic girdle were on average 0.8% larger than fish with a girdle; $t = 0.873$, $df = 8$, $P = 0.408$), nor among the surviving fish in both treatments combined (fish without a pelvic girdle were on average 0.9% larger than fish with a girdle; $t = 1.0374$, $df = 8$, $P = 0.3299$).

3.3. Growth rates experiment

Fish without a pelvic girdle had a slight tendency to grow faster than their siblings with a pelvic girdle (Figures 3 and 4; Table A.4) ($t = 1.459$, $df = 8$, $p = 0.183$). Across all families, the mean growth rate of fish without a pelvic girdle was 0.112 mm/day. Fish with a pelvic girdle grew 0.107 mm/day, around five percent less than fish without a pelvic girdle. There was considerable variation in growth rates within the families themselves, thus, the 95% confidence interval for the mean of the differences (*without*s - *with*s) was wide (approximately -0.003 to 0.014 mm/day). All fish had similar standard lengths, within one or two millimeters, at the start of the experiment; however, after 28 weeks, body sizes were widely distributed. In all families, a few ‘runts’ were present, small fish that did not grow and had high mortality during the experiment. The results did not change significantly when I analyzed only the larger half of the fish of each family at each measurement, thus removing the effect of the ‘runts’ (95% confidence interval for differences (*without*s - *with*s) from -0.003 to 0.015 mm/day; $t = 1.428$, $df = 8$, $P = 0.191$).

4. Discussion

4.1. Insect predation hypothesis

In predation experiment 1, the fish with the shortest pelvic spines tended to survive best. The estimate of the selection differential for pelvic spine length was comparable both in direction and strength to previous estimates from experiments on F_2 hybrid stickleback, which expressed high variation in this trait (Marchinko, 2009). In contrast to those experiments, however, which had measured selection on spine length in fish that also segregated many other traits, predation experiment 1 measured the direct effects of pelvic spines on survival in the face of predation by backswimmers. Pleiotropic effects of the genes underlying spine length were excluded, as well as any effects of correlated traits.

In insect predation experiment 2, fish without a pelvic girdle were 80% more likely to survive in the face of insect predation, but the standard error was large and the null hypothesis could not be rejected. It thus appeared that in this experiment either factors other than presence or absence of the pelvic girdle influenced the fishes' odds for survival, or the sample size was simply too small to confirm an effect. Fish without a pelvic girdle usually have no pelvic spines. Under the hypothesis that pelvic structures increase an individual fish's probability of being caught by insect predators, they were expected to have an advantage over fish with a girdle and spines. However, the lack of the pelvic girdle itself might be disadvantageous. Alternatively, since the fish were backcrosses, those individuals possessing a pelvic girdle might also have possessed other, disadvantageous traits or gene combinations tightly linked to the main girdle locus (*Pitx1*). Such linkage could be reduced by generating further-generation backcrosses.

Contrary to the expectation, there was no overall directional selection for larger body size in predation experiment 1. In predation experiment 2, however, a trend for selection for larger body size was present. One explanation for the lack of selection on body size in insect predation experiment 1 might be that large fish usually also had very long spines. While short spines appeared to be generally advantageous, very small fish with short spines might have been easy to catch due to

their small size, while large fish with very long spines (standard length was positively correlated with pelvic spine length in the fish that were used in this experiment; Table A.3) might have been more vulnerable than expected if size were the only predictor variable.

4.2. Growth rates hypothesis

The hypothesis that the absence of a pelvic girdle allowed juvenile fish to grow faster received weak support from the growth rate measurements. Fish with a pelvic girdle grew about five percent slower on average than their siblings without a pelvic girdle, but confidence in this result was low. To be demonstrated conclusively, an effect of this size would require a larger sample size than was achieved in this study. This does not mean, however, that a five percent faster growth is ecologically negligible, even less so as wild juvenile stickleback are probably under intense predation. For a population of stickleback inhabiting a lake where piscivorous fish were present, Reimchen (1994) estimated mortality before reproduction to lie between 97 and over 99.9%.

This experiment was carried out on crosses between fish lacking a pelvic girdle and fish with a pelvic girdle, from the same population (Paxton Lake benthics). The pelvic structures of fish in this population were size-reduced, compared to the ancestral marine pelvic girdle. In populations with larger and more robust pelvic structures, the growth rate difference between fish that lose their girdle through a mutation in the *Pitx1* locus and fish that retain their pelvic structures, might be larger than the effect measured in this study. One other aspect of the experimental conditions should be discussed briefly. The fish of the growth rates experiment were raised in very calcium poor water in an attempt to maximise the costs of having to build a pelvic girdle. However, the fish used in this experiment were descendants from a population inhabiting calcium rich water (Paxton Lake), and might therefore have had difficulties coping with the low calcium levels in their ambient water during the experiment. Also, it was not quantified whether phosphates and other essential minerals for building bony pelvic structures were available in sufficient quantities. Nutrient stress can result in slower overall growth (Arendt, 1997). If nutrient stress was indeed present, growth rates of all fish might have been equally reduced, and an effect of presence or absence of a pelvic

girdle might still have been measurable. Alternatively, it is conceivable that the response to nutrient stress varied among fish, as maximum attainable growth rates of sticklebacks are at least partly genetically determined (Wright et al., 2004). Unless such an effect was linked to the main pelvic girdle locus, it could have diminished growth differences between the two pelvic phenotypes. Unfortunately, the experiment did not have a control treatment for calcium limitation.

Over all, the presented data provided qualified support for the hypotheses of differential predation through backswimmers (*Notonecta*) on pelvic spine phenotypes, and to a lesser extent, pelvic girdle phenotypes. Also, growth might be slower in fish with a pelvic girdle. All observed effects were small (5 to 10%-differences between the phenotypes, but with large standard errors). The small sample sizes used in this study may obscure the significance of such small differences between the phenotypes.

4.3. Comparison with other data

Field studies on natural selection often do not achieve the power necessary to demonstrate the effects of interest (Hersch and Phillips, 2004). It is also likely that experimental studies on selection often fall short of obtaining the sample sizes necessary to detect weak to moderate levels of selection. In addition, usually only statistically significant results are published. This seems to preclude a majority of those field studies (and potentially experimental studies) on natural selection, that detect weak selection, from being published (Hersch and Phillips, 2004). However, publication bias can be avoided if all data is made publicly available, including non-significant results, together with an estimate of power (Kingsolver et al., 2001; Hersch and Phillips, 2004). Additionally, Hersch and Phillips suggest that new studies be designed to be more powerful and to analyze selection on at least several hundred individuals.

The experiments of my study fall into the category of low power studies. The measured effects showed trends in the predicted directions, but were relatively small, and the small sample sizes led to wide confidence intervals. It is not feasible to carry out such experiments with hundreds of

trials, as might be necessary to achieve sufficient power (Appendix B.4; Figure C.1). I therefore compared my results with previously published experiments on invertebrate predation on stickleback, which are also suffering from low power. A list of the studies and their attributes for these analyses can be found in Table 3; the methods are detailed in Appendix B.

4.3.1. Selection on pelvic spine length

Marchinko (2009) is the only other study to have measured selection on spine length of three-spine sticklebacks in the presence of insect predators. The study also measured selection on other morphological traits (dorsal spine length, standard length and pelvic girdle length). The two F_2 hybrid lines (*Paxton line* and *McKay line*) used in the two experiments were both freshwater - marine crosses, and showed variation in pelvic spine length, but also correlated variation between other traits (e.g. dorsal spines and standard length) and pelvic spine length. Covariance with standard length was controlled for by analyzing the residuals from a regression between the two traits (pelvic spine length and standard length). The experiment was designed to measure the effect of a chromosomal region, which included several armour traits, on survival of fish in the face of insect predation. Thus the effects of the individual morphological armour traits were not independent. Selection on residual pelvic spine length was reported as standardized selection differentials. This allowed me to obtain a combined weighted estimate of the effect size over the two experiments from Marchinko and my insect predation experiment 1, using methods of meta-analysis for fixed effects (Cooper and Hedges, 1994, Appendix B). The resulting mean standardized selection differential was -0.155, the measuring unit being phenotypic standard deviations of residual pelvic spine length. The 95% confidence interval ranged from -0.036 to -0.275. This implied weak to moderate selection for shorter pelvic spine length in the face of insect predation (Figure 5 and Table 3A).

4.3.2. Selection on standard length

The mean selection differential for standard length was obtained by combining the results from my insect predation experiments 1 and 2 with those of the two experiments in Marchinko (2009),

which were also published as standardized selection differentials (Figure 6 and Table 3B). The sizes of fish used in all four experiments ranged from 10 to 30 mm. The combined selection differential, (fixed effects model, for details see Appendix B) therefore, represented selection on juvenile fish, from appearance of spines (around 10 mm) to a size where there are largely invulnerable to most invertebrate predation (Foster et al., 1988). The combined standardized selection differential was positive (0.203) and although the 95% confidence interval was wide (0.045 to 0.360), it was consistent with the hypothesis that invertebrate predation imposed weak to moderate selection for greater body size in juvenile stickleback.

4.3.3. Selection on presence or absence of the pelvic girdle

Two previously published studies by Reist (1979) and Marchinko (2009) also tested experimentally whether presence of a pelvic girdle affected vulnerability of juvenile stickleback to insect predators. The results of these studies were combined with those of my insect predation experiment 2. Two additional studies (Reimchen, 1980; Ziuganov and Zotin, 1995) were only qualitatively analyzed. I calculated the log odds ratios of survival for fish lacking a pelvic girdle for each experiment (mixed effects model; for methods see Appendix B), and combined them. Unlike the multiple experiments on body size and pelvic spine length, the direction of the effects was not consistent among studies (Figure 7 and Table 3C). Two of the experiments, those by Reist and my predation experiment 2, displayed positive mean log odds ratios (with large standard errors), which suggested that a loss of the pelvic girdle led to higher survival in the face of insect predation. On the other hand, Marchinko's experiment, using F₂ hybrids between Paxton benthic (freshwater) and Oyster Lagoon (marine) threespine stickleback, demonstrated a negative log odds ratio. The weighted mean effect estimate of all experiments combined was a 4.4% decrease in survival probability for fish without a pelvic girdle in the presence of insect predation, compared to fish with a girdle. The 95% confidence interval was wide. At the lower limit, survival probability of fish *with* a pelvic girdle was 106.7% higher than that of fish without a pelvic girdle. At the upper limit, fish *without* a pelvic girdle had a 89.8% higher survival probability, compared to the other phenotype

(Figure 7 and Table 3C).

Compared to the other studies, which all reported a tendency for fish without a pelvic girdle to be at an advantage, the negative log odds ratio estimate of Marchinko (2009) was puzzling. Since it had a small standard error, it contributed strongly towards the weighted mean. Differences in experimental conditions as well as sampling error might be responsible for these divergent results. An alternative explanation for Marchinko's estimate is that the experiment was carried out on fish from a cross between two highly differentiated populations of *G. aculeatus*. Negative fitness epistasis has been suggested as a factor affecting the results, decreasing the fitness of certain genotypes of fish without a pelvic girdle in the predation experiment, through mechanisms that could not be addressed by the study (Marchinko, 2009).

The two studies that were not included in the meta-analysis (Reimchen, 1980; Ziuganov and Zotin, 1995) both suggested an increased survival probability for fish without a pelvic girdle, compared to fish with a pelvic girdle. From Reimchen (1980) I calculated a log odds ratio of 0.159, which stands for a 17% increase in survival probability for fish without pelvic structures. From the results reported in Ziuganov and Zotin (1995) I could not calculate a log odds ratio for survival. The study included an insect predation experiment on ninespine sticklebacks (*Pungitius pungitius*). The fish came from a population that displays pelvic girdle polymorphism, including complete loss of the pelvic girdle. In the face of insect predation, the study found a strong selective advantage for fish lacking the pelvic girdle, compared to fish with a complete girdle. Since the experiment had no replicate trials, confidence in this result could not be quantified.

The results from the combined analysis indicated that the direction of selection on presence or absence of the pelvic girdle through invertebrate predation cannot be conclusively determined from the available studies. Direction of selection might depend on the predators present, on the stickleback species under selection, and other, unknown factors. The net effect across all populations might be positive, neutral, or even negative selection with regard to loss of the pelvic girdle. In some respects, this result might reflect the situation in nature, where pelvic girdle loss is only found in a fraction of all extant freshwater stickleback populations, despite the fact that most of them are probably exposed to insect predation. Whether insect predation has a different impact in

each population, or whether its effect on pelvic girdle frequency is comparatively small and often drowned by other selective forces remains to be determined.

The mean effect size estimates that were obtained comparing previous experiments with my own data (Table 3), were comparable in magnitude to published estimates from other experimental studies on selection on phenotypic traits (Morjan and Rieseberg, 2004). Estimates of selection differentials from experimental studies are generally larger than estimates from natural populations (Morjan and Rieseberg, 2004). This is especially true for manipulative experiments like the spine clipping in predation experiment 1, which produced great variation in trait values. Interestingly, such high variation might also be found under natural circumstances during the beginning of freshwater habitat colonization by marine stickleback, when mutations in regulatory regions of the main pelvic girdle locus (*Pitx1*) might lead to a sudden loss of the pelvic girdle and spines. At this point, pelvic girdle loss could be driven predominantly by the disadvantage of having long and robust pelvic spines, even if the pelvic girdle itself, as a whole, were not disadvantageous.

4.4. Conclusions

Is predation by aquatic invertebrates a selective factor for pelvic girdle loss in freshwater populations of threespine stickleback? My predation experiment 1 suggested that insect predation might select for shorter pelvic spines in juvenile stickleback. Predation experiment 2 indicated that fish with a pelvic girdle might have reduced survival probability in the face of insect predation, compared to fish lacking pelvic structures, but confidence in this result was low. A meta-analysis, combining my results with previously published insect predation experiments on pelvic phenotypes, provided clearer answers. It showed that invertebrate predators selected for shorter pelvic spine length in juvenile stickleback. The pelvic girdle (with or without long or short pelvic spines) appeared to be disadvantageous in the face of insect predation in some experiments, but overall, no direct selection for loss of the pelvic girdle was observed. These findings offer support for the original hypothesis by Reimchen (1980), that pelvic spines make it easier for insect predators to hold on to their prey. Pelvic girdle loss might be a by-product of the selection for the loss of pelvic

spines, as these two morphological traits (pelvic spines and girdle) are usually linked. The meta-analysis also showed that smaller fish were generally more vulnerable to invertebrate predation than larger fish.

Do fish with a pelvic girdle grow more slowly than fish without a girdle? My growth rates experiment tested if there was a difference between growth rates of juvenile sticklebacks with and without a pelvic girdle. The experiment estimated, but failed to confirm, that growth rates of juveniles with a pelvic girdle were indeed five percent lower than those of fish without pelvic structures. Such an effect, if confirmed by further studies, combined with selection against small body size through insect predators, could be part of the mechanism behind pelvic girdle loss in freshwater threespine stickleback populations.

Table 1: Standardized selection differentials and significance tests (t -test) for insect predation experiment 1. Reported are the mean standardized selection differentials, standard error (SE), and the 95% confidence interval (CI). Selection differentials for the individual trials can be found in Appendix A (Tables A1-A3).

	Mean	SE	95% CI	$t_{0.05(2)}$	<i>df</i>	<i>P</i>
Pelvic spine length (residuals)	-0.163	0.073	-0.341 to 0.014	-2.252	6	0.065
Pelvic spine length	-0.176	0.096	-0.411 to 0.060	-1.827	6	0.118
Standard length	-0.049	0.132	-0.372 to 0.275	-0.367	6	0.726

Table 2: Comparisons among insect predation experiments on threespine and brook stickleback.

		Effect size measure	Mean	SE	95% CI lower	95% CI upper	Number of experiments
A	Pelvic spine length (residuals)	Selection differential	-0.155	0.061	-0.275	-0.036	3
	Marchinko - Paxton line		0.258	0.469			
	Marchinko -McKay line		-0.160	0.115			
	This study - Experiment 1		-0.163	0.073			
B	Standard length	Selection differential	0.203	0.080	0.045	0.360	4
	Marchinko -Paxton line		0.645	0.212			
	Marchinko -McKay line		0.225	0.140			
	This study - Experiment 1		-0.049	0.132			
	This study - Experiment 2		0.337	0.201			
C	Pelvic girdle presence/absence	Log odds ratio	-0.043	0.349	-0.726	0.641	3
	Reist experiments		0.209	0.605			
	Reimchen*		0.159	-			
	Marchinko -Paxton line		-0.462	0.234			
	This study - Experiment 2		0.587	0.542			

Effect size measures are either *standardized selection differentials* or *log odds ratios*. The weighted means, standard errors (SE) and 95% confidence intervals (CI) were calculated using meta-analysis methods (see Appendix B). In the beginning of each section (A, B and C) are the results from the meta-analyses, given below are the means and standard errors of the individual experiments included in the comparison. *Number of experiments* is the number of individual experiments analyzed in the meta-analyses. *Was not included in the weighted mean.

Table 3: Aquatic insect predation experiments on pelvic girdle phenotypes (pelvic girdle presence or absence) in *C. inconstans*, *G. aculeatus* and *P. pungitius*. *Results within trials not independent (see text). **Were not included in the weighted mean.

Author and year	Fish species	Predator species	Traits	Fish size [mm]	Replicate trials	Fish total	Comments
Reist 1979	<i>C. inconstans</i>	<i>Lethocerus americanus</i>	Pelvic girdle	20 - 29	11*	149	30 l tanks; natural population
	<i>C. inconstans</i>	<i>Dysticus spp.</i>	Pelvic girdle	20 - 29	7*	98	30 l tanks; natural population
	<i>C. inconstans</i>	<i>Dysticus spp.</i>	Pelvic girdle	20 - 29	5*	69	30 l tanks; natural population
	<i>C. inconstans</i>	<i>Aeshna spp.</i>	Pelvic girdle	20 - 29	7*	91	30 l tanks; natural population
Reimchen 1980**	<i>G. aculeatus</i>	<i>Aeshna spp.</i>	Pelvic girdle	15 - 25	7*	408	natural population
Marchinko 2009	<i>G. aculeatus</i> (Paxton line)	<i>Aeshna spp.</i>	Pelvic spines; pelvic girdle; standard length	10 - 18	6	477	380 l enclosures; F ₂ hybrids
	<i>G. aculeatus</i> (McKay line)	<i>Aeshna spp.</i> <i>Notonecta sp.</i>	Pelvic spines; standard length	10 - 23	10	767	380 l enclosures; F ₂ hybrids
This study	<i>G. aculeatus</i> (Experiment 1)	<i>Notonecta sp.</i>	Pelvic spines; standard length	11 - 22	7	423	189 l tanks; natural population
	<i>G. aculeatus</i> (Experiment 2)	<i>Notonecta sp.</i>	Pelvic spines; pelvic girdle; standard length	9- 18	9	573	380 l enclosures; backcrosses
Ziuganov & Zotin 1995**	<i>P. pungitius</i>	Odonata spp. <i>Dysticus spp.</i>	Pelvic girdle	51-62	1*	200	50'000 l pond (exp); 9'000'000 l pond (con); natural population

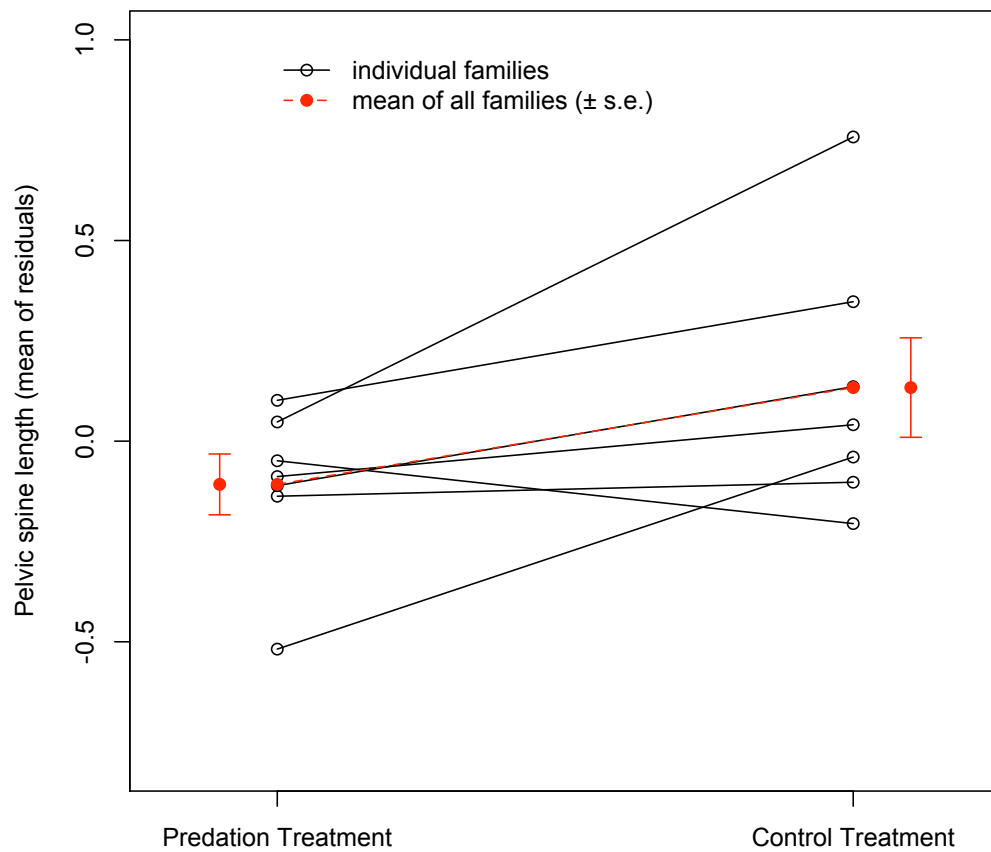


Figure 1: Insect predation experiment 1. Mean pelvic spine length (residuals) of surviving fish in the predation and control treatments. Also shown are means and standard errors of the treatments.

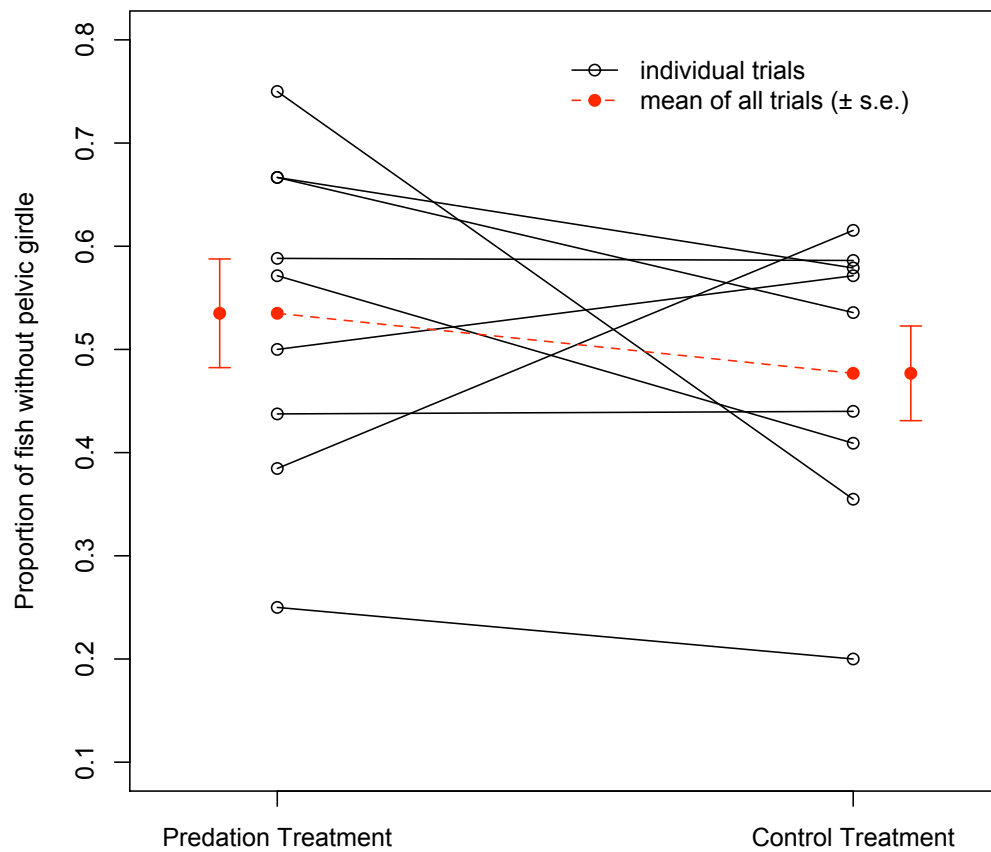


Figure 2: Insect predation experiment 2. Proportion of fish without a pelvic girdle among all fish that have survived the predation and the control treatment. Also shown are means and standard errors of the treatments.

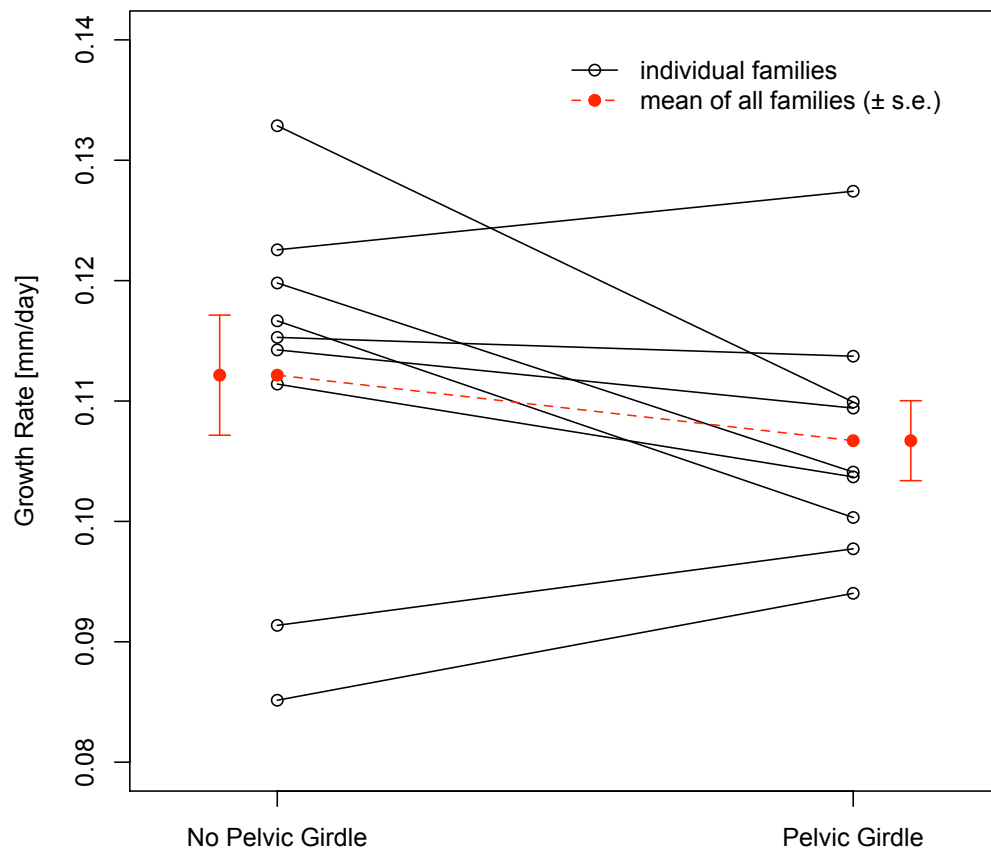


Figure 3: Growth rates of all families, including 100% of all fish alive at each measurement, comparing the two pelvic girdle phenotypes. Also shown are the means and standard errors of each phenotype..

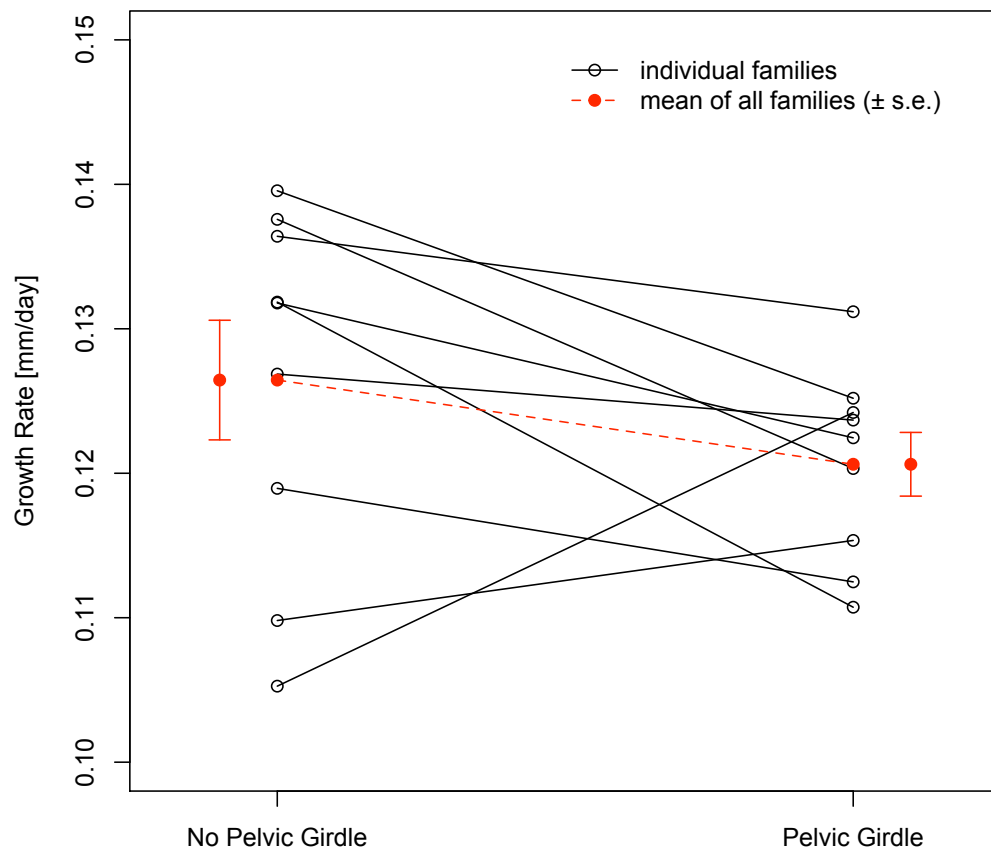


Figure 4: Growth rates of all families, only including the largest 50% of the fish at each measurement, comparing the two pelvic girdle phenotypes. Also shown are the means and standards error of each phenotype.

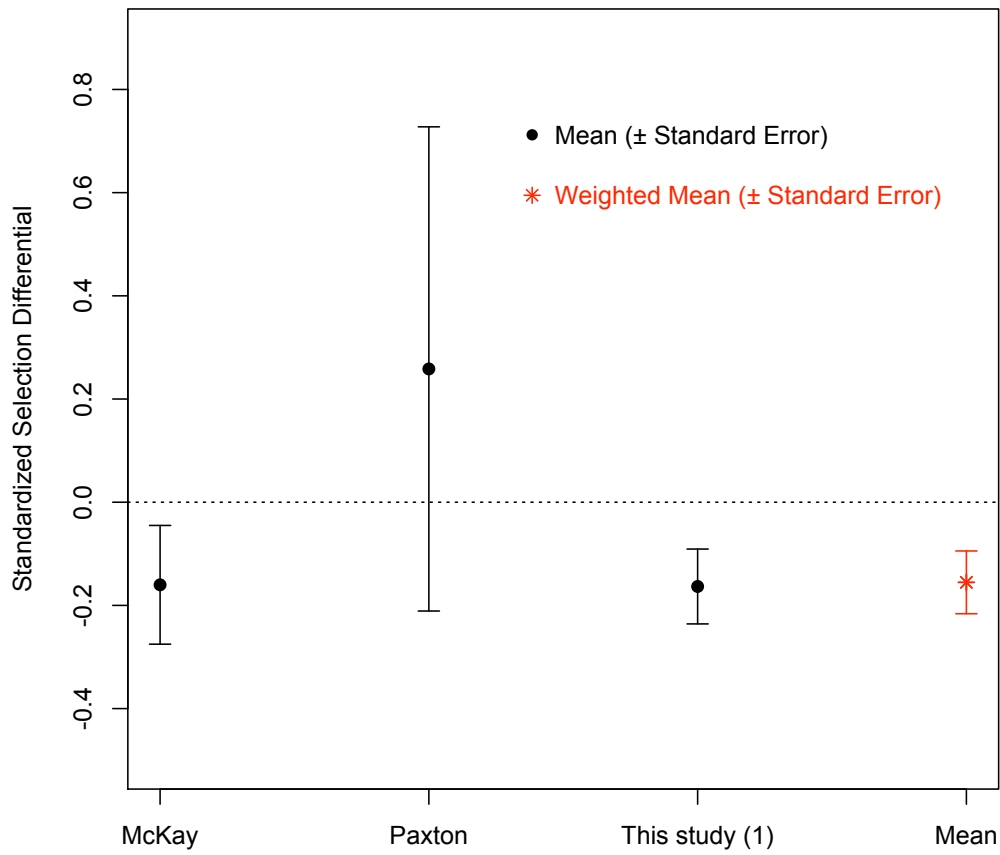


Figure 5: Mean standardized selection differentials for pelvic spine length. The across-experiments weighted mean was calculated by using the inverse variance method, applying a fixed effects model. A negative selection differential stands for increased survival as pelvic spine length decreases. *McKay* and *Paxton* refer to the two experiments in Marchinko (2009). *This study (1)* refers to my insect predation experiment 1.

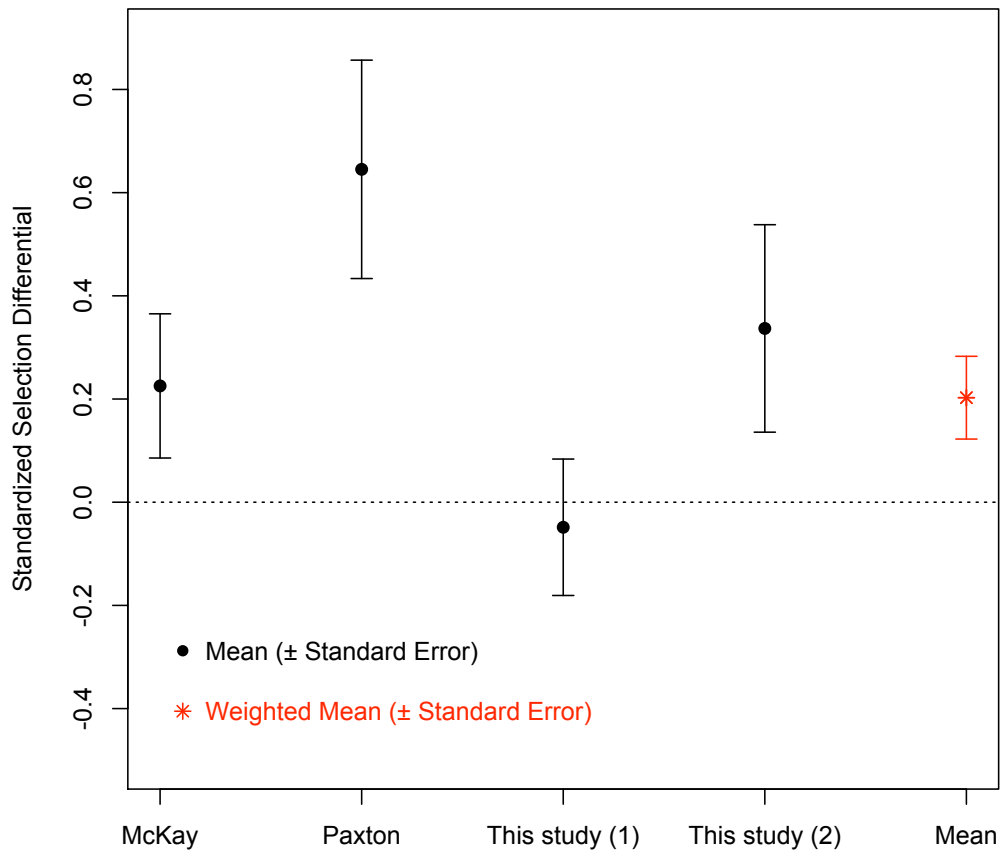


Figure 6: Mean standardized selection differentials for standard length. The across-experiments weighted mean was calculated by using the inverse variance method, applying a fixed effects model. A positive selection differential stands for increased survival with increased standard length of the fish. *McKay* and *Paxton* refer to the two experimental lines in Marchinko (2009). *This study (1)* and *(2)* refer to my insect predation experiments 1 and 2.

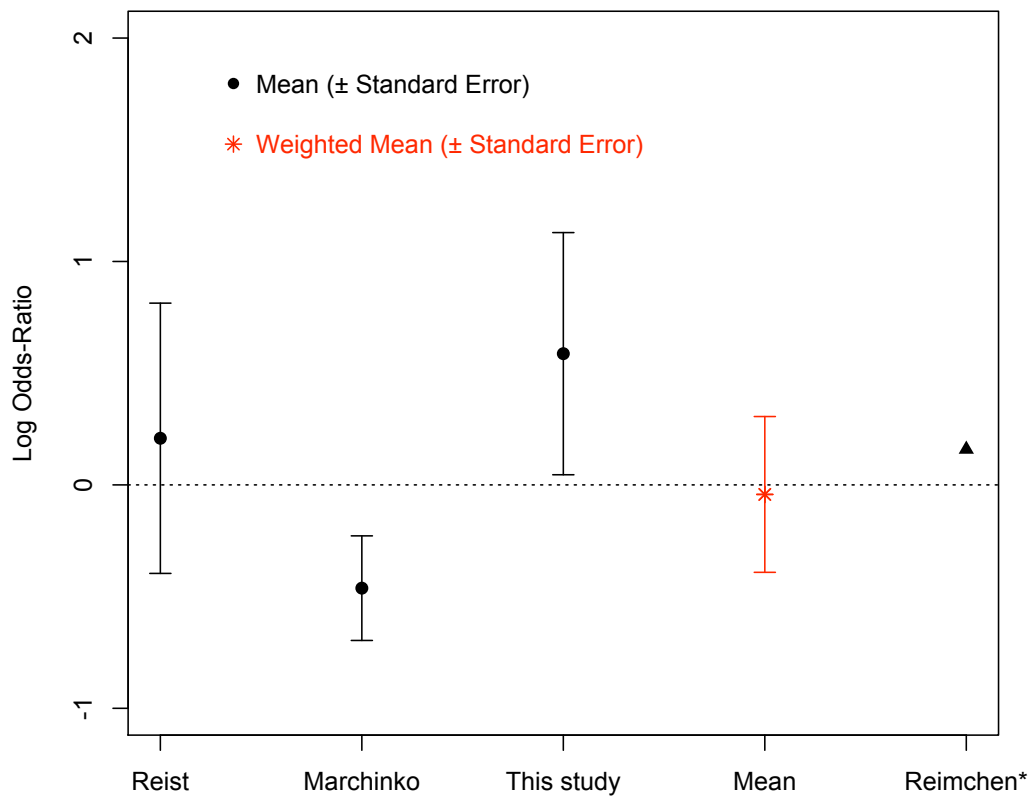


Figure 7: Mean log odds ratios from the different insect predation experiments. The across-experiments weighted mean was calculated by using the inverse variance method, applying a mixed effects model. A log odds ratio of zero equals no effect of the pelvic girdle on survival in the presence of aquatic insects. *No standard error, and thus not included in the weighted mean (see text).

References

- P.A. Abrams. The evolution of predator-prey interactions: theory and evidence. *Annual Review of Ecology, Evolution, and Systematics*, 31:79–105, 2000.
- N.E.W. Alstad, L. Skardal, and D.O. Hessen. The effect of calcium concentration on the calcification of *Daphnia magna*. *Limnology and Oceanography*, 44:2011–2017, 1999.
- J.D. Arendt. Adaptive intrinsic growth rates: an integration across taxa. *The Quarterly Review of Biology*, 72:149–177, 1997.
- R.D.H. Barrett, S.M. Rogers, and D. Schluter. Natural selection on a major armor gene in three-spine stickleback. *Science*, 322:255–257, 2008.
- J.V. Baumgartner, M.A. Bell, and P.H. Weinberg. Body form differences between the enos lake species pair of threespine sticklebacks (*Gasterosteus aculeatus* complex). *Canadian Journal of Zoology*, 66:467–474, 1988.
- M.A. Bell. Interacting evolutionary constraints in pelvic reduction of threespine sticklebacks, *Gasterosteus aculeatus* (pisces, gasterosteidae). *Biological Journal of the Linnean Society*, 31:347–382, 1987.
- M.A. Bell, G. Orti, J.A. Walker, and J.P. Koenings. Evolution of pelvic reduction in threespine stickleback fish: a test of competing hypotheses. *Evolution*, 47:906–914, 1993.
- M.A. Bell, M.P. Travis, and D.M. Blouw. Inferring natural selection in a fossil threespine stickleback. *Paleobiology*, 32:562–577, 2006.
- C.W. Benkman. Divergent selection drives the adaptive radiation of crossbills. *Evolution*, 57: 1176–1181, 2003.
- M. Borenstein, L.V. Hedges, Higgins J.P.T., and Rothstein H.R. *Introduction to Meta-Analysis*. John Wiley & Sons, Ltd, 2009.

- G.M. Clarke. Developmental stability and fitness: The evidence is not quite so clear. *The American Naturalist*, 152:762–766, 1998.
- H. Cooper and L.V. Hedges. *The Handbook of Research Synthesis*. Newbury Park, CA: Russell Sage Foundation, 1994.
- S.M. Coyle, F.A. Huntingford, and C.L. Peichel. Parallel evolution of *pitx1* underlies pelvic reduction in scottish three spine stickleback. *Journal of Heredity*, 98:581–586, 2007.
- W.A. Cresko, A. Amores, C. Wilson, J. Murphy, M. Currey, P. Phillips, M.A. Bell, C.B. Kimmel, and J.H. Postlethwait. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National Academy of Sciences*, 101: 6050–6055, 2004.
- R.A. Curry, S.L. Currie, S.K. Arndt, and A.T. Bielak. Winter survival of age-0 smallmouth bass, *Micropterus dolomieu*, in Northeastern lakes. *Environmental Biology of Fishes*, 72:111–122, 2005.
- J.A. Endler. *Natural Selection in the Wild*. Princeton Univ. Press, Princeton, NJ, 1986.
- S.A. Foster, V.B. Garcia, and M.Y. Town. Cannibalism as the cause of an ontogenetic shift in habitat use by fry of the threespine stickleback. *Oecologia*, 74:577–585, 1988.
- N. Giles. The possible role of environmental calcium levels during the evolution of phenotypic diversity in Outer Hebridean populations of three-spined stickleback, *Gasterosteus aculeatus*. *Journal of Zoology*, 199:535–544, 1983.
- D.W. Hagen and J.D. McPhail. The species problem within *Gasterosteus aculeatus* on the Pacific coast of North America. *Journal of the Fisheries Research Board of Canada*, 27:147–155, 1970.
- E.I. Hersch and P.C. Phillips. Power and potential bias in field studies of natural selection. *Evolution*, 58:479–485, 2004.

- R.D. Hoogland, D. Morris, and N. Tinbergen. The spines of sticklebacks (*Gasterosteus* and *Pygosteus*) as a means of defense against predators (*Perca* and *Esox*). *Behaviour*, 10:205–236, 1957.
- G. Hunt, M.A. Bell, and M.P. Travis. The evolution of mutation rates. *Evolution*, 62:700–710, 2008.
- J.G. Kingsolver, H.E. Hoekstra, J.M. Hoeskstra, D. Berrigan, S.N. Vignieri, C.E. Hill, A. Hoang, P. Gilbert, and P. Beerli. The strength of phenotypic selection in natural populations. *The American Naturalist*, 157:245–261, 2001.
- J.P. Koenings, J.A. Edmundson, G.B. Kyle, and J.M. Edmundson. *Limnology field and laboratory manual: methods for assessing aquatic production*. Alaska Department of Fish and Game. FRED Division Report Series. 71, 1987.
- R. Lande and S.J. Arnold. The measurement of selection on correlated characters. *Evolution*, 37: 1210–1226, 1983.
- R. B. Langerhans. Evolutionary consequences of predation: avoidance, escape, reproduction, and diversification. In A.M.T. Elewa, editor, *Predation in organisms: a distinct phenomenon.*, pages 177–220. Springer-Verlag, Heidelberg, Germany, 2006.
- G.L. Larson. Social behavior and feeding ability of two phenotypes of *Gasterosteus aculeatus* in relation to their spatial and trophic segregation in a temperate lake. *Canadian Journal of Zoology*, 54:107–121, 1976.
- C.M. Lessels and P.T. Boag. Unrepeatable repeatabilities: a common mistake. *The Auk*, 104: 116–121, 1987.
- K.B. Marchinko. Predation’s role in repeated phenotypic and genetic divergence of armor in three-spine stickleback. *Evolution*, 63(127–138), 2009.

- K.B. Marchinko and D. Schluter. Parallel evolution by correlated response: Lateral plate reduction in threespine stickleback. *Evolution*, 61:1084–1090, 2007.
- J.D. McPhail. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): evidence for a species-pair in Paxton Lake, Texada Island, British Columbia. *Canadian Journal of Zoology*, 70:361–369, 1992.
- J.R. Meyer and R. Kassen. The effects of competition and predation on diversification in a model adaptive radiation. *Nature*, 446:432–435, 2007.
- A.P. Møller. Developmental stability and fitness: a review. *The American Naturalist*, 149:916–932, 1997.
- C.L. Morjan and L.H. Rieseberg. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology*, 13:1341–1356, 2004.
- P. Nosil and B.J. Crespi. Experimental evidence that predation promotes divergence in adaptive radiation. *Proceedings of the National Academy of Sciences*, 103:9090–9095, 2006.
- H.A. Orr. The genetics of species differences. *Trends in Ecology and Evolution*, 16:343–350, 2001.
- H.A. Orr. The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, 6:119–127, 2005.
- C.L. Peichel, K.S. Nereng, K.A. Ohgi, B.L. Cole, P.F. Colosimo, C.A. Buerkle, D. Schluter, and D.M. Kingsley. The genetic architecture of divergence between threespine stickleback species. *Nature*, 414:901–905, 2001.
- R.H. Peterson and D.J. Martin-Robichaud. Perivitelline and vitelline potentials in teleost eggs as influenced by ambient ionic strength, natal salinity, and electrode electrolyte; and the influence of these potentials of cadmium dynamics within the egg. *Canadian Journal of Fisheries and Aquatic Sciences*, 43:1445–1450, 1986.

- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- T.E. Reimchen. Spine deficiency and polymorphism in a population of *Gasterosteus aculeatus* - an adaptation to predators. *Canadian Journal of Zoology*, 58:1232–1244, 1980.
- T.E. Reimchen. Structural relationship between spines and lateral plates in threespine stickleback (*Gasterosteus aculeatus*). *Evolution*, 37:931–946, 1983.
- T.E. Reimchen. Predators and morphological evolution in threespine stickleback. In M.A. Bell and S.A. Foster, editors, *Evolution of the threespine stickleback*, pages 240–273. Oxford University Press, 1994.
- T.E. Reimchen and C.A. Bergstrom. The ecology of asymmetry in stickleback defense structures. *Evolution*, 63:115–126, 2009.
- T.E. Reimchen and P. Nosil. Temporal variation in divergent selection on spine number in three-spine stickleback. *Evolution*, 56:2472–2483, 2002.
- J.D. Reist. Predation upon pelvic phenotypes of brook stickleback, *Culaea inconstans*, by selected invertebrates. *Canadian Journal of Zoology*, 58:1253–1258, 1979.
- D. Reznick and J.A. Endler. The impact of predation on life history evolution in trinidadian guppies (*Poecilia reticulata*). *Evolution*, 36:160–177, 1982.
- B.W. Robinson and S.L. Wardrop. Experimentally manipulated growth rate in threespine sticklebacks: assessing trade-offs with developmental stability. *Environmental Biology of Fishes*, 63: 67–78, 2002.
- F.J. Rohlf. *TpsDig2 software*. SUNY, Stony Brook, NY, 2006. URL <http://life.bio.sunysb.edu/morph/>.

- H.D. Rundle, S.M. Vamosi, and D. Schluter. Experimental test of predation's effect on divergent selection during character displacement in sticklebacks. *Proceedings of the National Academy of Sciences*, 100:14943–14948, 2003.
- D. Schluter. *The ecology of adaptive radiation*. Oxford University Press, Oxford, UK, 2000.
- E.T. Schultz, L.M. Clifton, and R.R. Warner. Energetic constraints and size-based tactics: the adaptive significance of breeding-schedule variation in a marine fish (embiotocidae: *Micrometrus minimus*). *The American Naturalist*, 138:1408–1430, 1991.
- M.D. Shapiro, M.E. Marks, C.L. Peichel, B.K. Blackman, K.S. Nereng, B. Jónsson, D. Schluter, and D.M. Kingsley. Genetic and developmental basis of evolutionary pelvic reduction in three-spine sticklebacks. *Nature*, 428:717–723, 2004.
- K. Simkiss. Calcium translocation by cells. *Endeavour*, 33:119–124, 1974.
- S.M. Vamosi. Predation sharpens the adaptive peaks: survival trade-offs in sympatric sticklebacks. *Annales Zoologici Fennici*, 39:237–248, 2002.
- S.M. Vamosi. The presence of other fish species affects speciation in threespine sticklebacks. *Evolutionary Ecology Research*, 5:717–730, 2003.
- S.M. Vamosi. On the role of enemies in divergence and diversification of prey: a review and synthesis. *Canadian Journal of Zoology*, 83:894–910, 2005.
- G.J. Vermeij. Unsuccessful predation and evolution. *The American Naturalist*, 120:701–720, 1982.
- M.C. Whitlock and D. Schluter. *The Analysis of Biological Data*. Roberts and Company, Greenwood Village, CO, 2009.
- H.A. Wright, R.J. Wootton, and I. Barber. Interpopulation variation in early growth of threespine sticklebacks (*Gasterosteus aculeatus*) under laboratory conditions. *Canadian Journal of Fisheries and Aquatic Sciences*, 61:1832–1838, 2004.

V.V. Ziuganov and A.A. Zotin. Pelvic girdle polymorphism and reproductive barriers in the nine-spine stickleback *Pungitius pungitius* (L.) from Northwest Russia. *Behaviour*, 132:1095–1105, 1995.

A. Appendix A

A.1. Mean pelvic spine length in predation experiment 1

Table A.1: Mean pelvic spine length [mm] for the predation and control treatments. SSD refers to the standardized selection differentials for pelvic spine length.

Trial	Mean	Mean	SSD
	Experiment	Control	
1	2.709	2.656	0.036
2	1.894	2.450	-0.434
3	3.082	3.955	-0.576
4	2.331	2.298	0.029
5	2.583	2.917	-0.203
6	2.492	2.737	-0.178
7	2.391	2.275	0.096

A.2. Mean standard length in predation experiment 1

Table A.2: Mean standard length [mm] for the predation and control treatments. SSD refers to the standardized selection differentials for standard length.

Trial	Mean	Mean	SSD
	Experiment	Control	
1	16.505	16.138	0.321
2	14.701	15.021	-0.248
3	17.287	17.962	-0.439
4	14.729	14.058	0.514
5	14.990	15.354	-0.311
6	15.494	15.486	0.005
7	14.817	14.986	-0.181

A.3. Mean pelvic spine length (residual) in predation experiment 1

Table A.3: Mean pelvic spine length (residuals) for the predation and control treatments. SSD refers to the standardized selection differentials for pelvic spine length (residuals).

Trial	Mean	Mean	SSD
	Experiment	Control	
1	-0.137	-0.102	-0.025
2	-0.518	-0.040	-0.352
3	0.048	0.758	-0.439
4	-0.088	0.041	-0.125
5	0.102	0.348	-0.153
6	-0.111	0.136	-0.182
7	-0.049	-0.205	0.134

Pelvic spine length was correlated with standard length of the fish. Combining all fish from all trials, of both clipping treatments, the Pearson's correlation coefficient was 0.269 (95% confidence interval from 0.165 to 0.367) ($t = 5.002$, $df = 321$, $P = 9.364e^{-7}$). The correlation coefficient for "tips-only"-clipped fish for all trials combined was 0.810 (95% CI from 0.749 to 0.857; $t = 17.365$, $df = 158$, $P \leq 2.2e^{-16}$). The correlation coefficient for "whole-spines"-clipped fish for all trials combined was 0.613 (95% CI from 0.507 to 0.701; $t = 9.850$, $df = 161$, $P \leq 2.2e^{-16}$). Residuals were calculated with an ordinary least square regression of spine length on standard length, combining the fish from all trials and treatments. The standardized selection differentials were calculated using the residuals.

A.4. Growth rates of the two pelvic girdle phenotypes

Table A.4: Growth rates for the two pelvic phenotypes of each family in the growth rate experiment, in [mm/day].

Family	Growth Rate	
	No Pelvic Girdle	Pelvic Girdle
1	0.094	0.095
2	0.118	0.098
3	0.082	0.090
4	0.109	0.108
5	0.106	0.098
6	0.087	0.091
7	0.107	0.094
8	0.129	0.106
9	0.123	0.131

Note: Families 1 and 2 shared the father, families 4, 5 and 6 shared the mother, and families 7 and 8 had the same father.

B. Appendix B

B.1. Methods for comparisons among experiments

I used methods of meta-analysis to compare the results of the available studies on selection on pelvic spine length, standard length and pelvic girdle. To calculate the averages of effects over studies I weighted the effect sizes of the different studies by the inverse of the sampling variance of the estimate of each study (Cooper and Hedges, 1994). The number of studies available was admittedly small, and the resulting confidence intervals for the average effect sizes were expected to be comparatively large.

B.2. Mean selection differentials

I determined mean selection differentials for pelvic spine length and standard length. Marchinko (2009) measured selection on pelvic spine length in the presence of insect predation on juvenile threespine stickleback in two separate experiments. Each experiment used an F_2 hybrid cross between one of two marine populations and one of two freshwater populations (two crosses in total) (data in Marchinko (2009): Supporting Information; Tables 2 and 3). The present study measured selection on spine length in one experiment (insect predation experiment 1), and selection on standard length in two experiments (insect predation experiments 1 and 2). For each of these four experiments, I calculated the mean selection differential (pelvic spine length and/or standard length) and the standard error. I then calculated the average selection differential for size and pelvic spine length using the *fixed effects model*, following Borenstein et al. (2009). The fixed effects model was deemed appropriate here because the experiments I compared were highly similar to each other, in each case involving a similar experimental design carried out on the same species (threespine stickleback) under similar laboratory conditions. Variation among the experiments in the estimate of selection was assumed to be the result of sampling error, rather than of inherent differences among the studies. The implication of using the fixed effects model is that the resulting effect size estimates should not be used to make inferences outside of the context of the analyzed

studies. In other words, the resulting mean selection differential for pelvic spine length is valid for experimental studies in which juvenile threespine stickleback exhibit large variation in pelvic spine length are exposed to predation by insect predators (*Aeshna spp.* and *Notonecta sp.*). The same implications hold for the mean selection differentials for standard length.

B.3. Mean log odds ratios

There are four published studies that have investigated pelvic girdle loss in the face of insect predation: Reist (1979), Reimchen (1980), Ziuganov and Zotin (1995) and Marchinko (2009). Details of these studies are summarized in Tables 2 and 3 (main text). In all experiments, juvenile or small adult stickleback (three different species) were exposed to a number of insect predators. The experiments were terminated when approximately half of the fish were dead, which was typically after 7 to 10 days (except in the experiment of Ziuganov and Zotin (1995), which lasted 55 days). The main difference among the experiments was that Marchinko (2009) used families of fish that were bred specifically for the experiment (F_2 hybrid crosses), while the other studies all used wild caught fish.

I combined the results of my insect predation experiment 2 with two of the published experiments on the effect of invertebrate predation on survival of different pelvic girdle phenotypes of juvenile stickleback. I used the log odds ratios (odds of surviving insect predation relative to the odds of dying) as the effect size of each experiment (Whitlock and Schluter, 2009). The experimental setup varied among the studies, requiring a different meta-analysis approach (see below). The trial is the appropriate unit of replication, which was the case in Marchinko's experiment and in my insect predation experiment 2, with each trial corresponding to a stickleback family. However, the two studies by Reist and Reimchen did not present results or analyses based on the number of trials. Instead, they presented results and analyses based on pooled data from all replicate predation trials (Tables B.1 and B.2). Thus I could not obtain standard errors of the effect sizes from the individual experiments in the Reist and Reimchen studies. To circumvent this problem the individual experiments of the Reist study ($n=4$) were treated as the units of replication. From the

log odds ratios of each of the four experiments of the study by Reist (1979), I calculated the mean log odds ratio and the standard error. The same was not possible for the Reimchen experiment, and no standard error could be obtained. The experiment was therefore not included in the meta-analysis. The fourth study, on the selective advantage of pelvic girdle phenotypes of *Pungitius pungitius* (Ninespine Stickleback), by Ziuganov and Zotin (1995), had no replication (only one treatment and one control pond were used) and thus it was not possible to calculate a standard error for their estimate of effect size. As a result it was not included in the analysis.

From the data of the Marchinko (2009) and insect predation experiment 2, I calculated the mean log odds ratio and standard error from the log odds ratios of the individual trials. Both experiments had two pelvic girdle categories: *No pelvic girdle*, which included all fish that either had no pelvic structures at all, or exhibited tiny, remnant pelvic structures only (see Bell et al., 1993). *With pelvic girdle* included all other fish, mostly exhibiting a complete pelvic girdle. In both experiments, the fish had not been scored (pelvic girdle presence/absence) or measured before the experiments started. The exact numbers of fish in each of the two pelvic girdle categories that had entered the experiment were therefore not known. In order to calculate odds ratios, it was necessary to know how many fish of each category had entered the predation treatment. Thus I estimated the starting ratios of fish with to fish without a pelvic girdle in the following way: I assumed that mortality in the control treatments had been random, and that therefore the frequencies of the pelvic phenotypes in the control treatments (after the experiment) were representative for the family as a whole (before the experiment). For each trial, I multiplied the ratio of the pelvic girdle categories in the control treatment (after the experiment) with the total number of fish that had entered the predation treatment (before the experiment). This gave me an estimate for the frequencies of the two pelvic categories in the predation treatment (before the experiment). I then calculated the log odds ratio (odds of surviving insect predation relative to the odds of dying) for the predation treatment of each trial, following Whitlock and Schluter (2009).

In contrast to the experiments measuring selection on pelvic spine length and standard length, the three pelvic girdle experiments were highly heterogeneous in experimental design and the type of predators and the species of stickleback used. As a result, the studies were expected to vary not

only due to sampling error, but also because each experiment was likely to be estimating a different effect size. For this reason, I used a mixed effects model to calculate an average log odds ratio for the three experiments on presence and absence of the pelvic girdle (again following Borenstein et al. (2009)). The mixed effects model differs from the fixed effects model in that it calculates a between-study variance in addition to the within-study variance. The between-study variance, however, is difficult to estimate precisely when the number of studies involved is small, as it is in my case ($n=3$). There is no satisfying solution to this problem (see Cooper and Hedges (1994), pp. 316-317; and Borenstein et al. (2009), p. 84), so I provided the estimate and confidence interval from the meta-analysis, but also refer to the underlying data (Table 2C, Figure 7).

I reported the mean effect sizes, standard errors and 95% confidence intervals as calculated using the respective appropriate model (fixed or mixed effect model) in sections 4.3.1 to 4.3.3. All analyses were carried out using the statistical software package R (R Development Core Team, 2009).

B.4. Power of pelvic girdle predation experiments

How many trials are needed to demonstrate an effect of the pelvic girdle on survival in the face of insect predation? I carried out a simple simulation of predation experiment 2. I assumed that each trial was exposing a family of 36 fish, with equal numbers of fish with and without a pelvic girdle (or any other trait), to a predator. Fish with a pelvic girdle had a certain advantage (10, 15, 20 or 30%) over the other phenotype. The experiment stopped when half of the fish were dead, and the log odds ratio of survival for fish lacking a pelvic girdle was calculated. I then determined the power of a given number of trials to detect (t -test, at the 0.05 significance level) a given advantage of fish without a pelvic girdle over fish with a girdle (Figure C.1).

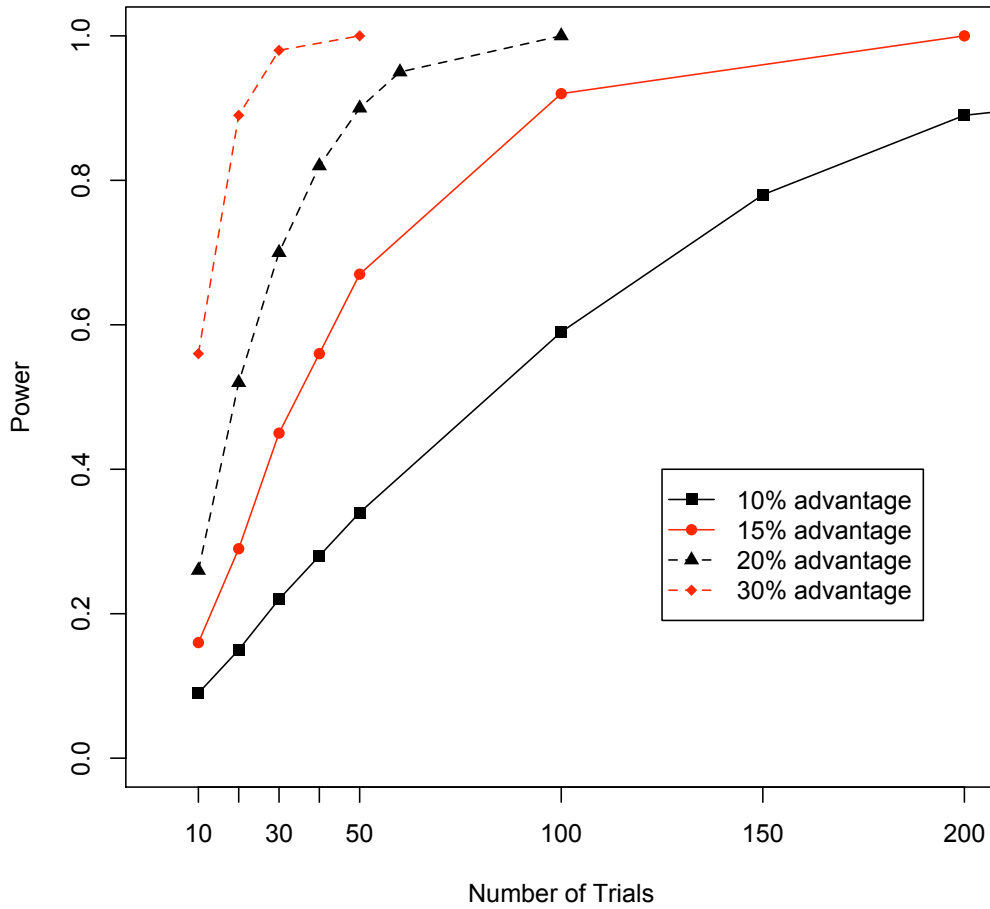


Figure B.1: Simulations of insect predation experiments that follow the design of predation experiment 2. Number of fish in each trial is 36, with equal numbers of the two phenotypes. The focal phenotype has a certain advantage (10, 15, 20 or 30%) over the other phenotype to not get killed during an attack by a predator. Each trial runs until 50% of all fish are dead. I estimated the power of a given number of trials to detect a positive effect of the focal phenotype (in this case fish without a pelvic girdle) onto survival (using log odds ratios as the effect size), at the 0.05 significance level (t -test). The power estimates are based on 1'000 to 10'000 runs of the simulation for each set of parameters.

Table B.1: Results of the experiments from Reist (1979). Each experiment reported numbers of individual fish, with the results of all trials pooled. The experiments were done on *Culaea inconstans*. On the top left of each table is the insect predator species, in brackets, where applicable, the dorsal spine phenotype of the fish used in the experiment.

<i>Lethocerus americanus</i>	Pelvic girdle	No pelvic girdle	Total
Not eaten	42	33	75
Eaten	34	40	74
Total	76	73	149

<i>Dysticus spp.</i> (5 dorsal spines)	Pelvic girdle	No pelvic girdle	Total
Not eaten	20	32	52
Eaten	29	17	46
Total	49	49	98

<i>Dysticus spp.</i> (6 dorsal spines)	Pelvic girdle	No pelvic girdle	Total
Not eaten	14	13	27
Eaten	21	21	42
Total	35	34	69

<i>Aeshna spp.</i>	Pelvic girdle	No pelvic girdle	Total
Not eaten	24	28	52
Eaten	21	18	39
Total	45	46	91

Table B.2: Results of the predation experiment on pelvic phenotypes of *G. aculeatus*, from Reimchen (1980). On the top left of the table is the insect predator species. The table represents the pooled numbers of individual fish from the seven trials that were conducted.

<i>Aeshna spp.</i>	Pelvic girdle	No pelvic girdle	Total
Not eaten	62	111	173
Eaten	93	142	235
Total	155	253	408

C. Appendix C

C.1. Animal care certificate

The Animal Care Committee of the University of British Columbia gave permission to carry out the experiments presented in this thesis (Figure C1).



ANIMAL CARE CERTIFICATE

Application Number: A07-0293

Investigator or Course Director: [Dolph Schluter](#)

Department: Zoology

Animals:

Sticklebacks Threespine stickleback 4000
Invertebrates Backswimmers (Notonecta) and dragonfly nymphs (Aeshna) 750
Trout Cutthroat trout 10
Fish Prickly sculpin 100

Start Date: April 1, 2006

Approval Date: November 17, 2008

Funding Sources:

Funding Agency: Canada Foundation for Innovation
Funding Title: CFI Infrastructure Operating Funds - The Origin and Persistence of Species - Operations

Funding Agency: Natural Sciences and Engineering Research Council of Canada (NSERC)
Funding Title: The genetics of adaptation to new environments

Funding Agency: Natural Sciences and Engineering Research Council of Canada (NSERC)
Funding Title: Ecology and genetics of adaptive radiation

Funding Agency: Canada Foundation for Innovation
Funding Title: CFI Infrastructure Operating Funds - The Origin and Persistence of Species - Operations

Funding Agency: Natural Sciences and Engineering Research Council of Canada (NSERC)
Funding Title: Ecology and genetics of adaptive radiation

Funding Agency: Natural Sciences and Engineering Research Council of Canada (NSERC)
Funding Title: The genetics of adaptation to new environments

Funding Agency: National Institutes of Health
Funding Title: 1 F32 GM086125-01 Genetics of Behavioral Reproductive Isolation in Threespine Stickleback, Schluter/Arnegard

Unfunded title: Ecology and genetics of adaptive radiation

The Animal Care Committee has examined and approved the use of animals for the above experimental project.

This certificate is valid for one year from the above start or approval date (whichever is later) provided there is no change in the experimental procedures. Annual review is required by the CCAC and some granting agencies.

A copy of this certificate must be displayed in your animal facility.

Office of Research Services and Administration

Figure C.1: Animal care certificate