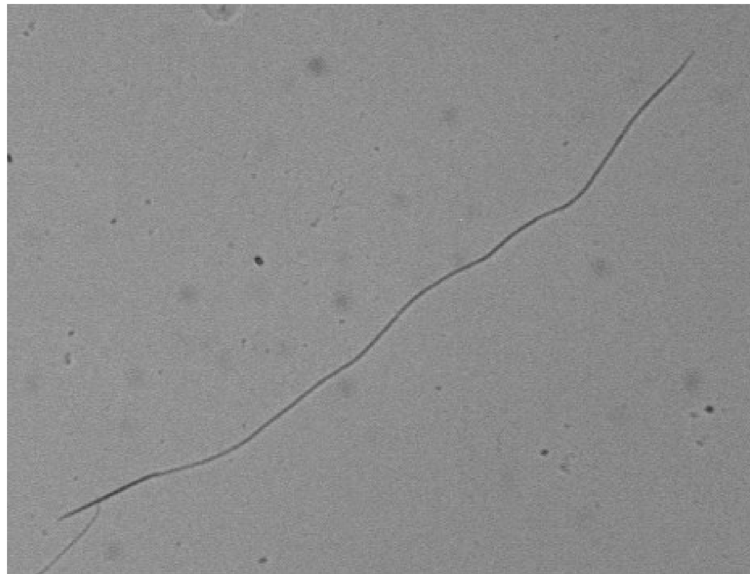
Are Longer Sperm Costly to Produce?

An Investigation with *Scathophaga stercoraria*

Diploma Thesis of
Ralph Dobler



Supervised by
Dr. David J. Hosken

Assessed by
Prof. Dr. Paul I. Ward

January 2005
Zoological Museum of the University of Zurich, Switzerland

Table of Contents

1. Abstract **3**

2. Introduction **4**

3. Material and Methods **8**

3.1 Sperm Length Selection **8**

Field Flies

Parental Lab Generations

Subsequent Generations

3.2 Assessing Potential Costs of Producing Longer Sperm **11**

Developmental Time

Longevity

Age at Sexual Maturity

Fertility

Immune System

3.3 Correlated Responses in Males **14**

Body Size

Testis Volume

3.4 Correlated Responses in Females **15**

Body Size

Developmental Time

Egg Size

Reproduction Organs

Immune System

3.5 Statistical Analyses **17**

4. Results 18

4.1 Sperm Length Selection 18

4.2 Assessing Potential Costs of Longer Sperm 21

Developmental Time

Longevity

Age at Sexual Maturity

Fertility

Immune System

4.3 Correlated Responses in Males 31

Body Size

Testis Volume

4.4 Correlated Responses in Females 33

Body Size

Developmental Time

Egg Size

Reproduction Organs

Immune System

5. Discussion 40

6. Acknowledgements 45

7. References 46

1. Abstract

Sperm length has considerable variation in nature. Sperm competition and female choice are two possible agents for this variation. In some species sperm length is huge and it is not clear how that could happen. As with every exaggerated trait, there should also be costs associated with producing long sperm. Trade-offs with other reproductive or life history traits are expected. To investigate potential costs of producing longer sperm I used bidirectional selection on sperm length in the yellow dung fly *Scathophaga stercoraria*. In addition I examined correlated responses to the selection in males and females. Sperm length significantly diverged with only four generations of selection, and I found, that males pay costs for longer sperm with increased developmental time and reduced fertility when mated with wild females. Moreover I found a significant body size x testis volume interaction between the two treatments. And there was evidence for a positive response of body size on sperm length selection. Females of L-lines had also an increased developmental time and tend to have longer spermathecal ducts. Like in *Drosophila spp.* there seem to be some costs related to production of longer sperm in the yellow dung fly *S. stercoraria*.

2. Introduction

Although sperm typically have a single function, to fertilise eggs, considerable variation is found in sperm form and size (e.g. Baccetti & Dallai 1978; Gage 1998; Hosken 2003; Oppliger *et al*. 1998, Ward & Hauschteck-Jungen 1993). Sperm form varies from spherical to "classical" sperm shape to amoeboid (e.g. Morrow & Gage 2001). Sperm size also varies from microscopic (28 micrometers in *Hystrix africaeaustralis*, Gage 1998), to relatively huge (5.8 centimetres in *Drosophila bifurca*, Pitnick *et al*. 1995).

Sperm competition is likely to be one selective agent responsible for the large variation in sperm length, and depending on the exact circumstances, sperm competition could favour larger or smaller sperm (Parker 1993). It has also been suggested that exaggerated sperm length is analogous to the peacock tails, and has resulted from female choice (Miller & Pitnick 2002). Evidence for this conjecture has been found in Drosophilids (Miller & Pitnick 2002), the genus showing the greatest sperm length variation (Pitnick *et al.* 1995).

As with any exaggerated trait, there must be costs to producing the long sperm found in many Drosophilids. A delay in male maturity has been reported in this genus for species with long sperm (Pitnick *et al.* 1995). Males from species with short sperm reach maturity before the females of the same species whereas males with long sperm reach maturity later than the females of their own species (Pitnick *et* al. 1995). The most likely reason for that delay may be, that in Drosophilids males need testes that are longer as the sperm they produce (Lindsley & Tokuyasu 1980). Therefore the observed delay of maturity may be a cost of building up and maintaining the sperm manufacturing machinery rather than the sperm production itself (Pitnick *et al.* 1995). How the time to maturity is impacted by sperm length is not known in species with less exaggerated sperm length. In addition, Pitnick and Miller (2000) found a similar effect on developmental time in *Drosophila hydei*. They found, that males with longer testes (and hence longer sperm) emerged later than males with shorter testes (and hence shorter sperm).

A trade-off between sperm size and number could also generate another potential cost to producing longer sperm (Gomendio & Roldan 1991). Within a limited energy budget producing longer sperm may mean that males have fewer of them (Parker 1982). This can reduce male fitness in two ways. Firstly males with longer sperm can be faced with sperm depletion when they mate with several females in a short time. Their fitness could consequently be reduced because the females they mated with may run out of sperm to fertilise their eggs. Secondly the sperm of the male could also be numerically disadvantaged during sperm competition. This will be the case when the female also mated with a male that had shorter (and therefore more) sperm or a male that is not yet in the state of sperm depletion. This assumes that the resources allocated to reproduction are fixed and hence trade-offs do not occur with non-reproductive traits. There are reports of sperm size versus sperm number trade-offs across species (e.g. Pitnick 1996, Pitnick & Markow 1994a).

However, trade-offs for reproductive traits do not have to be restricted to other reproductive traits. One of the key trade-offs in evolution occurs between reproduction and longevity (e.g. Stearns 1992). This trade-off may occur because reproduction is more important than a long life. Hosken (2001) showed that reproductive traits can trade-off with the immune system in the yellow dung fly *Scathophaga stercoraria*. He found that flies from polyandrous lines had lower phenyleoxidase (PO) activity level than flies from monogamous lines. PO activity level is an important part of the insect immune system (Söderhäll & Cerenius 1998) which can be used to assess the quality of the immune system. Polyandrous lines had larger testes than monogamous lines, the trade-off was between testes growth and investment to the immune system.

To investigate the potential costs of long sperm production, I used bidirectional selection on sperm length to create lines of flies (*S. stercoraria*) that significantly differed in sperm length. As potential costs which can occur when males produce longer sperm the following were assessed: developmental time, longevity, age at maturity, fertility across matings and PO activity level (to assess the quality of the immune system). Additionally, a number of characters in males and females that may have undergone correlated evolution as a result of the selection on sperm length were investigated. Pitnick (1996) found that males with longer sperm had larger testis and larger body size in *Drosophila spp*. I thus measured body size and testis volume in the selected males,

furthermore I also looked at body size and developmental time in females as these characters responded to similar selection (testis length) in *Drosophila hydei* (Pitnick & Miller 2000). Possible correlated response in egg size due to selection on male sperm length was also investigated, because both (sperm and egg) are gametes and may therefore be under similar genetic control. Additionally, I looked for possible correlated responses in female reproductive organs. The organs I investigated were the spermatheca, the spermathecal ducts and the accessory glands. Evidences for co-evolution between sperm length and some of these female traits have been found in different species (e.g. Miller & Pitnick 2002 for *Drosophila melanogaster*; Minder *et al.* in press for Scathophagidae). Miller and Pitnick (2002) showed that sperm length is an important factor for fertilisation. They found that males with longer sperm had a better fertilisation performance than males with shorter sperm, when the female they were mated with had a large seminal receptacle (sperm storage organ). In Scathophagidae sperm length and spermathecal duct length are positively correlated at species level (Minder *et al*. in press). Finally I had also a look at the PO activity level of the females (to assess the quality of the immune system).

I used the yellow dung fly *S. stercoraria* as model organism, and subjected replicate lines to four generations of bidirectional selection on sperm length. This species is an excellent model organism for bidirectional sperm length selection, because sperm length is highly heritable (Ward 1998, 2000) and not related to body size (Hosken *et al.* 2001; Ward & Hauschteck-Jungen 1993). Moreover variation in sperm length between individuals is greater than within an individual (Ward & Hauschteck-Jungen 1993). This means that different males from the same population produce sperm of different mean length. In *S. stercoraria* testis size, but not sperm length evolved when sperm competition levels were experimentally varied during laboratory evolution (Hosken *et al.* 2001). The monogamous lab rearing of these normally polyandrous flies should therefore have no impact on sperm length. However sperm length is influenced by environmental conditions, e.g. temperature (Blanckenhorn & Hellriegel 2002, Hellriegel & Blanckenhorn 2002). The change from the (fluctuating) field temperature to the constant lab temperature may therefore have an influence on sperm length. But this environmental change should have the same effect on all flies.

In my diploma thesis I tried to find answers to the following two specifically asked questions: 1) are there detectable costs to the production of longer sperm in *S. stercoraria*, and 2) are there correlated responses to selection on sperm length in males and / or females?

3. Material & Methods

3.1 Sperm length selection

Field Flies

Pairs of wild *S. stercoraria* (n=66) were caught on a pasture in Fehraltorf near Zurich (Switzerland) in mid-October 2003 and taken to the laboratory. In the laboratory each pair was placed into a glass vial and given a portion of fresh cow dung on a filter paper in which females could lay their eggs. Females were allowed to lay eggs over night, and 57 females laid eggs. Females from the field had (normally) already copulated with other males before they were brought to the lab. Therefore the (genetic and therefore the sperm length) variability in the parental lab generation was bigger than expected when females were only mated with one male.

Parental Lab Generation

About 15 to 20 eggs per female from the field flies were used for the parental lab generation and reared under standard conditions. Larvae were always reared under the same standard environmental conditions: 20° Celsius, 66% relative humidity and a 12:12 dark light regime. They were held in 100 ml plastic bottles filled with approximately 75 ml cow dung, this ensured *ad libitum* larval food conditions (more than 2g dung per larva (Amano 1983)) were established.

After 20 days offspring from the field flies began to emerge. All bottles were checked daily and the number of emerging flies was recorded (sex and family). Collected flies were always housed singly in a 100 ml glass bottle and provided with *ad libitum* water, sugar and excess *D. melanogaster* as prey under standard environmental conditions, except now (parental lab generation) flies were maintained at 13° Celsius (fore some logistical reasons). To establish the selection lines, males and females that emerged around the mean emergence time (males: 25.896 ± 0.248 days; females: 24.592 ± 0.123 days) were randomly divided into six lines. Three lines (replicates) were assigned to

the short sperm selection treatment, the other three, to the long sperm selection treatment. Each line consisted of about 40 to 45 males and the same number of females. After being assigned to their respective lines, flies were re-housed and fed as above. All flies were stored under standard conditions (except the rearing temperature was only 13° Celsius) until matings took place.

Matings took place on three successive days (one line per treatment per day). Females were at least 14 days old and males at least 6 days old, this ensured, that all flies were sexually mature (e.g. Hosken *et al.* 2002 for females; Foster 1967 for males). In each line 35 pairs were provided with the opportunity to copulate. Males were first added to a vial with a portion of cow dung on a filter paper. Females were added 10 to 25 minutes later. As always, matings between siblings were avoided. After a full copulation (longer than 15 minutes: direct observation), males were removed (after they released the female voluntarily) and stored in Eppendorf tubes at -20° Celsius for subsequent sperm length measurement. Females were allowed to lay one clutch of eggs over the next four days, although most laid within 24 hours. About 20 eggs per female were then transferred to a 100 ml plastic bottle filled three-quarter with cow dung and stored under standard conditions until offspring emergence.

Sperm length was measured for all frozen males and additionally, the hind tibia length (HTL), which is a good reference for body size (Sigurjónsdóttir 1980; Ward and Simmons 1991), was also measured. To measure sperm length, one testis per individual was dissected out in a droplet of distilled water on a microscope slide. All parts other than the testis (fat droplets, tracheae and other tissue) were removed. The testis was pierced in the proximal third (close to the ejaculatory duct) to obtain only mature sperm. The testis was then removed and the sperm gently diluted in the droplet of distilled water and dispersed on the slide. After the slide was air-dried, sperm length could be measured using microscope images (x 160) conveyed to a PC running ZEISS KS300 software. 15 sperm per individual were measured. Preliminary investigation suggested this was sufficient to obtain an accurate average of a male's mean sperm length.

Mean sperm length was then calculated for each line from the blind measured males. The offspring of males with shorter sperm than the mean in the short sperm selection treatment (or longer sperm

than the mean in the long sperm selection treatment respectively) were kept for the next generation. When less than eleven families per line fulfilled this condition, the families from fathers with the next shortest sperm (i.e. slightly longer sperm than the mean in the short sperm selection treatment or slightly shorter sperm than the mean in the long sperm selection treatment) were taken to ensure eleven families were used per line in the next generation. This was necessary in five lines in generation one (S2 two families, S3 three families, L1 three families, L2 one family and L3 one family) and two lines in generation two (S2 and S3 two families each).

Subsequent Generations

As in the parental generation, flies were held singly in 100 ml glass bottles under standard conditions after emergence. Flies for matings were chosen at random within each line (to avoid a direct selection on developmental time) and than mated (avoiding sib-pairs). Again males were stored in Eppendorf tubes and frozen at -20° Celsius after they had a full copulation. About 20 eggs per female were transferred to approximately 75 ml cow dung in a 100 ml plastic bottle as in the parental generation. Sperm length and HTL were measured for each male and line mean was calculated as above.

Males with shorter (or longer, depending on treatment) sperm were then chosen and their offspring used to found the subsequent generation (as above). This continued for further three generations (equals four generations of selection in all) at which time I investigated potential costs of producing long sperm and looked for correlated responses to the sperm length selection. I also calculated the realised heritability of sperm length using two methods: response and weighted selection differentials (Falconer & Mackay 1996), and a linear regression approach (Falconer & Mackay 1996). Calculating the realised heritability was not part of the initial experimental aims. As a result, the calculations at best provided a rough estimate of the realised heritability of sperm length as I did not include unselected control lines in my treatments.

3.2 Assessing potential costs of producing longer sperm

Developmental time

Means of developmental time of males were calculated for each line (and hence treatment) during the collection of emerging flies at each selection generation. HTL was reported (see above) and used as covariate in the analyses.

Longevity

Two different experiments investigated potential survival costs associated with the production of longer sperm. Potential survival costs may occur at different times during the life of male flies. After emergence males with longer sperm may need more energy to maintain the production of longer sperm than males with shorter sperm after emergence (e.g. Foster 1967). To test this hypothesis the first experiment was conducted (details see below). Males with longer sperm may also invest more energy in producing longer sperm before emergence. Here the second experiment (details see below) could give us an answer. First and second experiments were established at the same time and were done with males from generation five (first unselected generation). Sperm length was still significantly different between the two treatments (ANOVA, sperm length as the dependent, treatment as the independent; $F_{1,10} = 72.415$; $P < 0.001$; Means ± SE: S-line males 213.479 ± 0.581 m, L-line males 220.623 ± 0.605 m). In the first experiment 15 males per line were held under standard conditions for ten days. Flies were then transferred into new vials and supplied only with *ad libitum* water (starvation conditions) and stored under standard environmental conditions. Twice a day (about 8 am and 6 pm) survival of the males was checked, date and time of death and HTL were recorded. In the second experiment, 15 males per line were held under standard conditions for three days before food deprivation. The rest of this experiment was identical to the first one.

Age at sexual maturity

To assess whether producing longer sperm delayed time to maturity, five flies per line were frozen singly in Eppendorf tubes from first day of emergence until the seventh day post-emergence (all males from generation five, first unselected generation), when testis size seemed to have reached a maximum. The maximal testis size can be used as an indicator for sexual maturity (Foster 1967). Developmental time was reported for each fly individually and used as covariate in the analyses. Flies had been reared under standard conditions. Males were dissected as for the sperm length measurement. HTL of each individual was recorded. Testes length and width were measured manually using binocular images (x 25) conveyed to a PC running ZEISS KS300 software.

Fertility

To assess the potential fertility costs of produce longer sperm, ten males per line each copulated consecutively with five females from the same line (this was to avoid any incompatibilities that may have arisen when copulations occurred across or outside of co-evolving males and females in a line) in generation five (first unselected generation). Again sib-matings were avoided. All copulation durations were recorded and used as a covariate in subsequent analyses. Copulation duration can be related to the number of transferred sperm during the copulation, and the number of transferred sperm can have an impact on fertilisation (Parker & Simmons 1994, Simmons *et al.* 1994). After a successful copulation males were transferred to new vials, and females were housed with a portion of cow dung for egg laying. When copulation was not successful (no copulation or shorter than 15 minutes), no subsequent copulation took place and these males (and the females they copulated with) were not included in any analyses. There were five males in S1, four males in S2, six males in S3, one male in L1, four males in L2 and four males in L3 which did not reach five complete copulations, leaving a sample size of five males in S1, six males in S2, four males in S3, nine males in L1, six males in L2 and six males in L3.

As soon as the female finished egg laying, about ten eggs were transferred to a moist filter paper with a tiny portion of dung. The total number of laid eggs was also recorded. Total number of laid

eggs was included as covariate in the analyses. Filter papers with dung and eggs were stored in Petri dishes under standard environment conditions. After 40 hours the number of unhatched eggs was counted and the fertility of the males was calculated as the hatch rate per clutch (based on the ten eggs). The first and the fifth female of each male got a new vial (only if first and fifth copulation were completed) and were housed again under standard conditions. Each week these females were again given the opportunity to lay eggs. Egg number and hatch rates were assessed as above. This procedure was repeated for as long as females lived to get the total number of offspring a male had with the first and the fifth female. Males and other females (than first and fifth) were frozen singly in Eppendorf tubes for further morphological measurements. Only data from males with five successful copulations were used for statistical analyses.

An additional ten males per line had the possibility to mate with a wild female. Data from this experiment were used to assess different fertility success of S-line males and L-line males with wild females in case females within lines had evolved to differ their sperm use efficiency for example. After the successful copulation the male was frozen at -20° Celsius in an Eppendorf tube for further determinations. After laying eggs, the female was transferred to a new vial and stored under standard conditions. Approximately ten eggs were transferred to a humid filter paper and treated as the eggs above, furthermore the total number of laid eggs was counted. After 40 hours the number of hatched larvae was checked and recorded to calculate the hatch rate (as above).

3.3 Correlated responses in males

Body size

To assess whether male body size evolved as a correlated response to sperm length selection, male HTL values from the experimental generation (generation five, first unselected generation) were used for statistical analyses.

Testis volume

To look for differences in the testes volume between the two treatments, mature males (generation five, first unselected generation) were dissected in the same way as described in the "Age at sexual maturity" experiment to measure testis length and width. Testis length and width were again measured using the ZEISS KS300 software running on a PC. A microscope mounted camera conveyed the picture to the PC. Testis volume was calculated with the formula for a rotation ellipsoid

$$V = \pi \cdot \frac{4}{3} \cdot \frac{l}{2} \cdot \left(\frac{w}{2}\right)^2$$

where l was the length and w the width of the testis.

3.4 Correlated responses in females

Body size

Females HTL were measured at the experimental generation (generation five, first unselected generation) to assess whether female body size gave a correlated response to selection in male sperm length.

Developmental time

Means of developmental time of females were calculated for each line (and hence treatment) during the collection of emerging flies at each selection generation. Unfortunately for females HTL were not measured. As a result, the mean male HTL was used in the analyses for males and females. This is a reasonable procedure because although males are bigger than females, the HTL of males and females within families is highly significantly correlated (Hosken *et al*. 2003). Since males and females within lines all come from the same families, this approximation seems justified.

Egg size

Eggs from ten females from each treatment were used to calculate the egg volume. Females were all from the experimental generation (generation five, first unselected generation). Eggs were measured with a binocular on 6.4 times magnification. Egg length and width were again measured using the ZEISS KS300 software running on a PC. A binocular mounted camera conveyed the picture to the PC. Egg volume was calculated with the formula for a rotation ellipsoid

$$V = \pi \cdot \frac{4}{3} \cdot \frac{l}{2} \cdot \left(\frac{w}{2}\right)^2$$

where l was the length and w the width of the egg.

Reproduction organs

Females two, three and four from the first part of the fertility experiment (generation five, first unselected generation) were used to assess potential correlated morphological responses (spermatheca (sperm storage organ) area, spermathecal duct (duct between *Bursa copulatrix* and the spermatheca) length and accessory gland (may release a fluid to lubricate during the copulation, see Hosken and Ward 1999) area (calculated with the formula for the ellipse area

$$A = \pi \cdot \frac{l}{2} \cdot \frac{w}{2}$$

with the length l and the width w)) in the two selection treatments. HTL was also measured to get an estimation of correlated response in body size. The females were dissected in a droplet of insect Ringers solution on a microscope slide. This way any artificial changes in the organs due to osmosis should be avoided. Spermatheca areas were measured automatically using binocular pictures (x 25) conveyed to a PC running ZEISS KS300 software. Spermathecal duct length was measured manually in the same way. Accessory gland length and width were measured manually at a 16 times magnification (see Minder *et al.* in press, for full descriptions of methods).

Immune system

Correlated responses of the immune system due to selection on sperm length was investigated by measuring the phenyleoxidase (PO) level of sexually mature flies. An insects major immune defence against large, multicellular parasites is the formation of melanised capsule around the foreign invader (Pathak 1993). PO is a key enzyme leading to this melanised capsule. About 15 virgin males and females of each line in generation five were stored at -80° Celsius after they reached maturity under standard conditions. Level of the PO activity was measured using the method of Schwarzenbach *et al.* (in press). Age and body size were included as covariates in the analysis.

3.5 Statistical Analyses

Statistical analyses were generally made with SPSS 11.0 for PC. Other methods are mentioned in the text where they were used. Data was checked for normal distribution and transformed if necessary.

When results were expected in a predicted direction based on findings from previous studies, one-tailed tests were employed. This was the case in developmental time (Pitnick & Miller 2000), age to sexual maturity (Pitnick *et al.* 1995), fertility (Miller & Pitnick 2002) and spermathecal duct length (Minder *et al*. in press).

4. Results

4.1 Sperm length selection

In each generation I used ANOVA and post-hoc FISHER'S LSD-tests with sperm length as the dependent variable and line as the factor to look at the effect of selection on sperm length. By generation four, results indicated that all three lines from the short sperm selection (S1 to S3) had sperm that was significantly shorter than the lines from the long sperm treatment (L1 to L3) ($F_{1,161}$ = 14.544; $P < 0.001$; mean sperm length ± SE [µm]: S1 211.43 ± 1.03; S2 215.81 ± 0.73; S3 213.66 ± 0.96; L1 219.69 ± 1.18; L2 218.60 ± 0.82; L3 220.19 ± 0.68; Figure 1; Table 1).

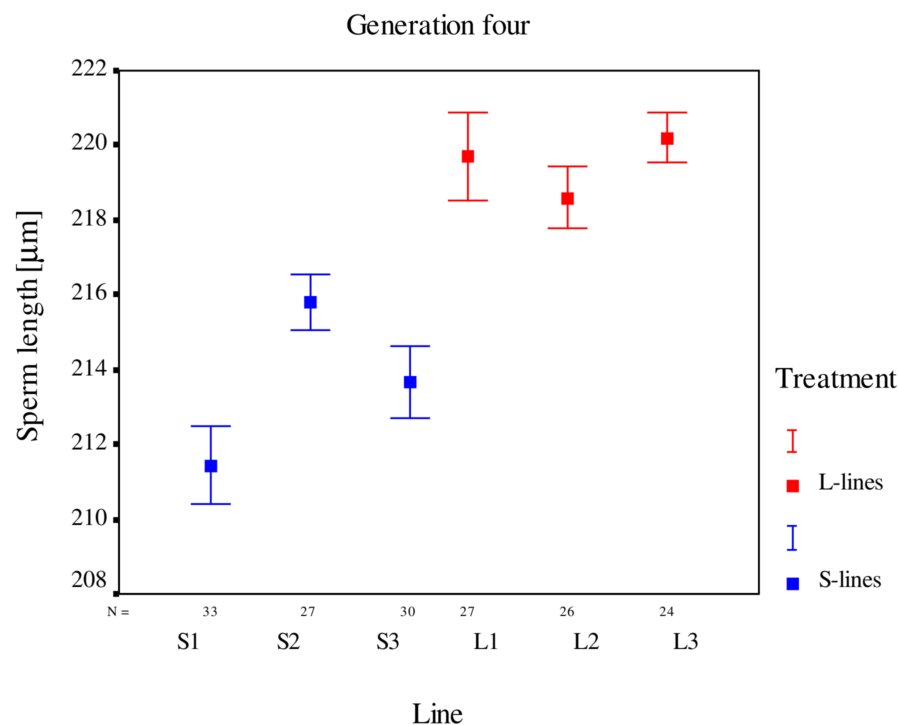


Figure 1: Mean (± SE) sperm length [µm] of each line at selection generation four. Generation four was the first generation when all three S-lines (blue) had shorter sperm than all three L-lines (red).

Flies reached a mean sperm length of 213.489 ± 0.570 µm in the S-lines by generation four and a mean of 219.483 ± 0.542 µm in the L-lines by generation four (Figure 2). As I started with a mean sperm length of 212.079 ± 0.389 µm for the S-lines and 212.284 ± 0.424 µm for the L-lines in the

first generation, results indicate that selection for longer sperm had a better response than selection for shorter sperm.

Table 1: P-values of the Fisher's LSD-test for the six selection lines at generation four. All S-lines (blue) are significantly different from all L-lines (red).

Line	S1	S2	S3	L1	L2	L3
S1	--	0.001	0.079	**<0.001**	**<0.001**	**<0.001**
S2		--	0.106	**0.016**	**0.044**	**0.002**
S3			--	**<0.001**	**<0.001**	**<0.001**
L1				--	0.426	0.721
L2					--	0.260
L3						--

Based on the response to selection over the four generations, I calculated the realised heritability for each treatment over the time of selection (four generations). I used response and weighted selection differential to calculate the means for each line. In S-line males realised heritability had a mean of

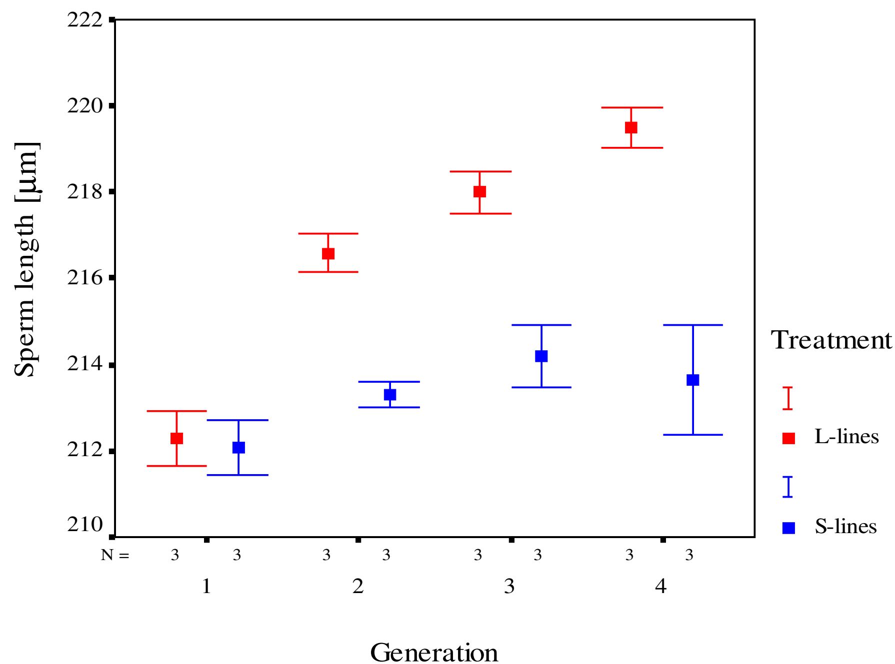


Figure 2: Mean (± SE) sperm length [µm] of each treatment (blue S-lines, red L-lines) over the four generations of selection.

0.089 ± 0.276 (SE), in L-line males 1.126 ± 0.366 (SE). Using linear regression I got a realised heritability of 0.312 ± 0.188 (SE) in the downward direction and 0.939 ± 0.143 (SE) in the upward.

4.2 Assessing potential costs of producing longer sperm

Developmental time

ANOVA was used to look at the effects of treatment (factor) on the developmental time (dependent) of males in generation four (when sperm length had significantly diverged). HTL was used as a covariate. Results indicated no significant treatment effect on developmental time ($F_{1,3} < 0.001$; $P = 0.495$; one-tailed test). HTL was not significant as covariate ($F_{1,3} = 3.702$; $P = 0.150$). S-line males had a mean developmental time of 25.778 ± 0.376 days by generation four and the L-line males had a mean developmental time of 26.630 ± 0.549 days by generation four (Figure 3). Results changed slightly when the non-significant covariate was removed ($F_{1,4} = 1.370$; $P = 0.154$; one-tailed test) but remained still non-significant, even though the results were in the predicted direction (S-line males slightly tended to have a shorter developmental time than L-line males).

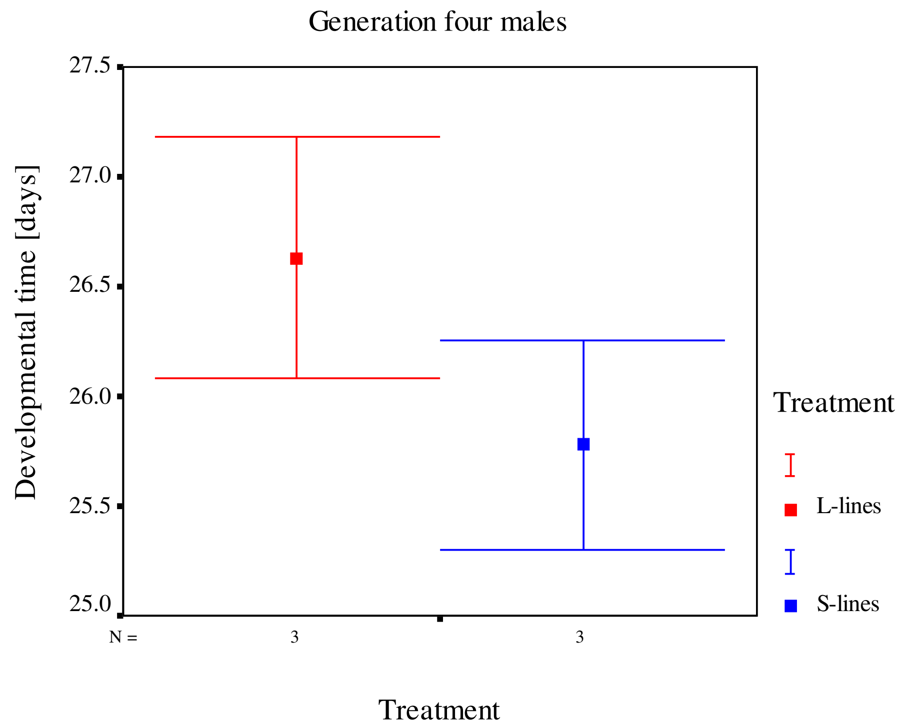


Figure 3: Mean (± SE) developmental time [days] for males of both treatments (blue S-lines, red L-lines) at generation four.

To further assess whether there may nevertheless be a difference in the developmental time of the two treatments an additional analysis was used to test these data. It employed linear regression techniques and compared the slopes of the two linear regressions (split by lines: i.e. one for S-lines and one for L-lines) of developmental time (Y-axis) against generation (X-axis), as these make use of the data from three rather than a single generation. With data from three generations the analysis had more statistical power. Slopes were compared using the methods of Zar (1998). Generation two had longer developmental time due to different rearing temperature (Blanckenhorn 1997). For this reason generation two was not directly comparable to the other generations and was excluded from the analyses. The slopes of these regression were in the predicted direction ($slope_S = 0.073$; $slope_L = 0.175$; developmental time of L-lines tend to become longer) and the differences were statistically significant ($|t| = 1.949 > t_{0.05(1),14}$; $0.05 > P > 0.025$; H_0: $slope_S = slope_L$ had to be rejected) (Figure 4).

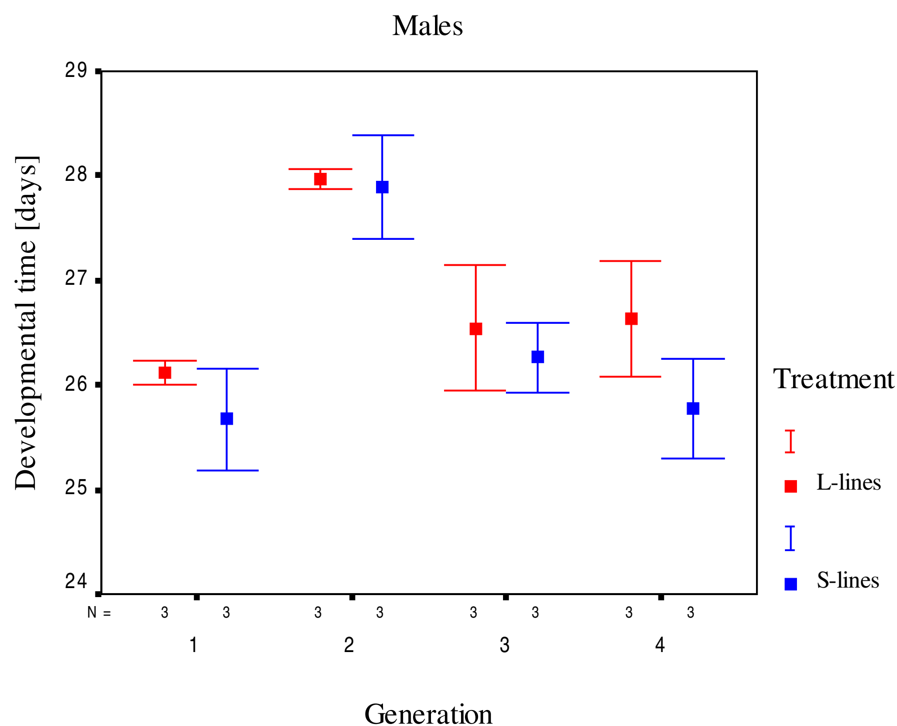


Figure 4: Mean (± SE) developmental time [days] for males of both treatments (blue S-lines, red L-lines) over the four generation with selection. Generation two was excluded from the analyses because rearing temperature was only 13° Celsius whereas all other generation were reared at 20° Celsius.

Longevity

ANCOVA was used to look at the effects of treatment (factor) on male longevity (dependent), and I used HTL and developmental time as a the covariates. Results indicated that treatment had no significant effect on the longevity ($F_{1,2} = 0.570$; $P = 0.529$) when flies were fed for ten days before starvation (Figure 5). S-line males lived for 24.511 ± 0.585 days, L-line males for 25.038 ± 0.693 days (Figure 5). When flies were fed for only three days before starvation, again treatment had no significant effect on the longevity ($F_{1,2} = 3.253$; $P = 0.213$; see also Figure 5). S-line males lived for 11.707±0.324 days and L-line males for 12.677 ± 0.491 days (Figure 5). HTL was not significant in either cases (ten day experiment: $F_{1,2} = 0.617$; $P = 0.515$; three day experiment: $F_{1,2} = 1.915$; $P = 0.301$) nor was developmental time (ten day experiment: $F_{1,2} = 0.692$; $P = 0.493$; three day experiment: $F_{1,2} = 0.460$; $P = 0.567$). When HTL and developmental time were removed as covariates (because they were not significant) conclusions did not change. No longevity costs could be found for males producing longer sperm assuming either maintenance or production costs. If anything, results indicate the opposite.

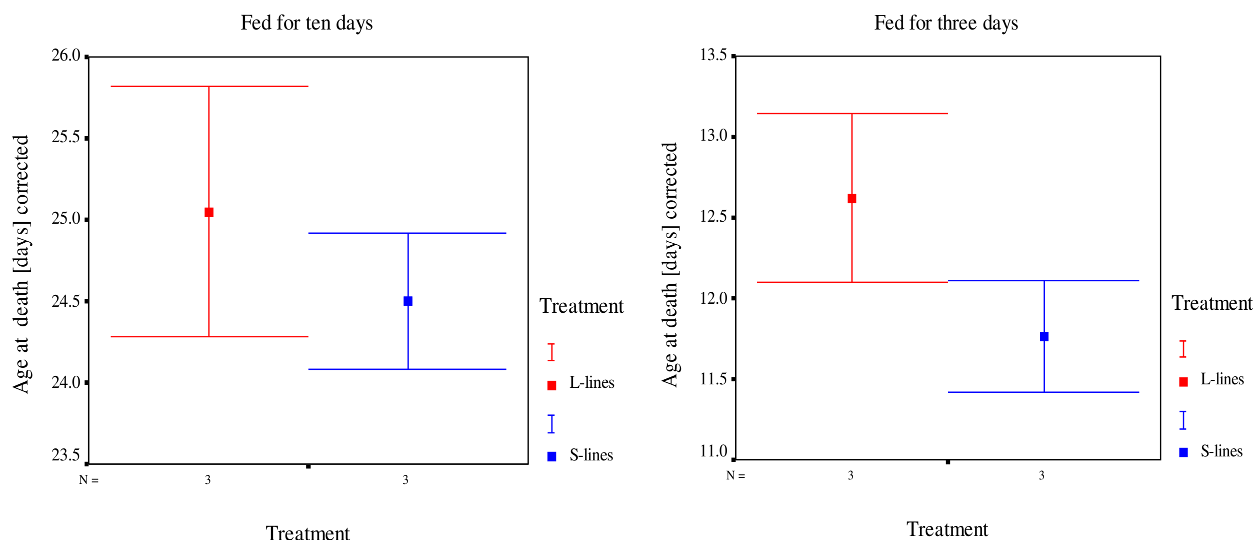


Figure 5: Average (± SE) life time of males (blue S-lines, red L-lines) in the two longevity experiments (left: Flies were fed for ten days before starvation, treatments were not different (P = 0.529); right: Flies were fed for three days before starvation, treatments were not different (P = 0.213).

Age at sexual maturity

I used ANCOVA to examine the effects of time after emergence and treatment (factors) on testis volume [(testis volume)$^{0.33}$ µm] (dependent) (as indicator of sexual maturity, see Foster 1967). HTL was used as the covariate. I found no significant age after emergence x treatment effect on testis volume ($F_{1,27} = 1.062$; $P = 0.205$; one-tailed test). Testis growth behave the same for both treatments in the first seven days post emergence. Results also indicated that age by itself had a significant effect on testis volume ($F_{1,27} = 2.506$; $P = 0.024$; one-tailed test; older males had larger testis volume), but treatment showed not more than a weak trend ($F_{1,27} = 0.863$; $P = 0.181$; one-tailed test). HTL as covariate was significant ($F_{1,27} = 5.250$; $P = 0.030$). As expected had larger males larger testis (e.g. Ward & Simmons 1991) (Figure 6).

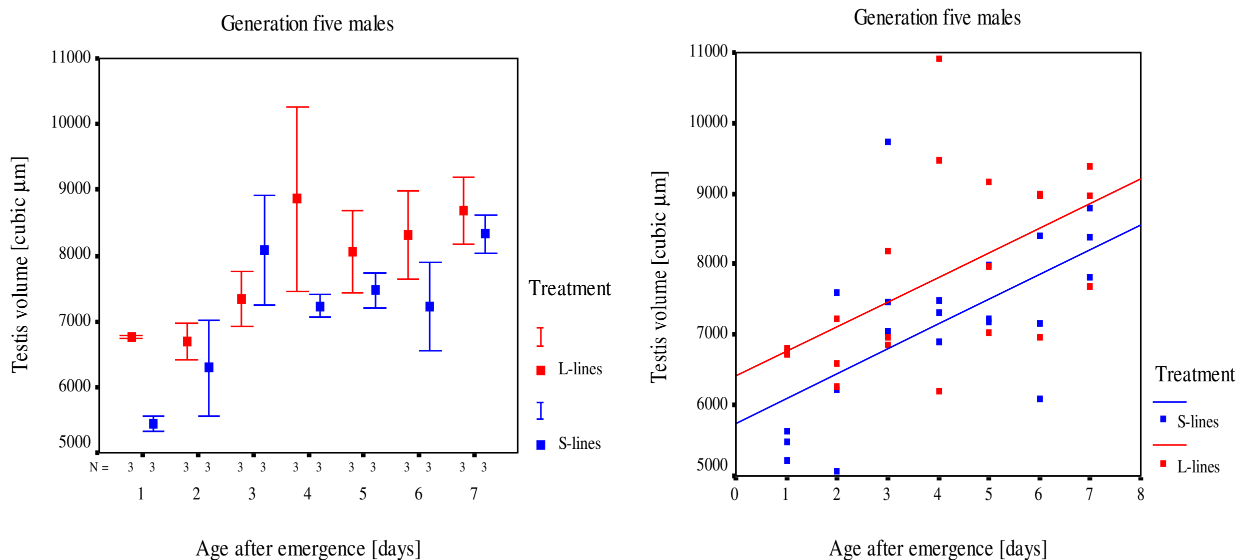


Figure 6: Left: Mean (± SE) testis volume [µm³] of the two treatment males (blue S-Lines, red L-lines) at each day post emergence. Right: Linear regression for testis volume increase over time (age after emergence) for the two treatments. $Slope_S$=0.164 and $slope_L$= 0.259, the two slopes were significantly different ($|t| = 1.958 > t_{0.05(1),42}$; $0.05 > P > 0.02$).

Two additional analyses were employed to further test these data. Both employed linear regression techniques. In the first of these the slopes of the two linear regressions (split by lines: i.e. one for S-lines and one for L-lines) of testis volume (Y-axis) against time after emergence (X-axis) were compared using the methods of Zar (1998). This way differences in the growth rate could be

assessed. The slopes of these regressions were 0.164 for the S-line regression and 0.259 for the L-line regression. Analysis indicated there was significant difference between the two. H_0: $slope_S = slope_L$ had to be rejected ($|t| = 1.958 > t_{0.05(1),42}$; $0.05 > P > 0.025$). This indicated that males from the L-lines had a faster testis growth rate than males form the S-lines (Figure 6). The second analysis used the methods of Yaeger and Ultsch (1989), to see if the inflection points (and hence the age at maturity) of the two regressions of the untransformed data differed. This analysis indicated that inflection was at the same time in both treatments. The inflection was between day three and four post emergence. This suggests that males from both treatments reached maturity at the same age post emergence and get the weak trend found in the ANCOVA out of the way.

Interestingly males from the S-lines had a significant smaller testis volume (5443.257 ± 120.323 μm^3) then males from the L-lines (6763.098 ± 29.025 μm^3) at day one after emergence (ANCOVA, testis volume [(testis volume)$^{0.33}$ µm] as the dependent, treatment as the fixed factor, HTL as the covariate, for each day post emergence; day one $F_{1,3} = 25.340$; $P = 0.015$; days two to seven $F_{1,3} < 1.440$; $P > 0.316$, see also Table 2). S-line males spend more energy in testis growth on day one after emergence and have therefore to compensate for this extra effort in the following days. This may explain the reduced longevity of S-line males, although it must be stressed that the longevity difference was not statistically significant.

Table 2: Effect of treatment on testis volume for each day post emergence. HTL was used as covariate in the analyses to correct for a possible effect of body size on testis volume.

	df	Mean Square	F	P
Day 1 after emergence				
HTL	1	8.824×10^{-4}	0.023	0.890
Treatment	1	0.989	25.340	**0.015**
Error	3	3.903×10^{-2}		
Day 2 after emergence				
HTL	1	2.725	9.968	0.051
Treatment	1	0.394	1.440	0.316
Error	3	0.273		

Table 2 continued

Day 3 after emergence				
HTL	1	0.209	0.172	0.706
Treatment	1	0.533	0.438	0.555
Error	3	1.215		
Day 4 after emergence				
HTL	1	0.571	0.231	0.664
Treatment	1	2.472	0.999	0.391
Error	3	2.474		
Day 5 after emergence				
HTL	1	4.744×10^{-2}	0.076	0.801
Treatment	1	3.616×10^{-2}	0.058	0.826
Error	3	0.627		
Day 6 after emergence				
HTL	1	1.434	1.713	0.282
Treatment	1	0.184	0.220	0.671
Error	3	0.837		
Day 7 after emergence				
HTL	1	0.102	0.247	0.653
Treatment	1	2.158×10^{-2}	0.052	0.834
Error	3	0.411		

Fertility

Several analyses were conducted to assess this possibility. Initially a repeated measurement ANCOVA was used to examine the effect of treatment on the number of hatched eggs [arcsine square root transformed proportion of hatched eggs] in the first or fifth copulation (repeated measure). Here treatment was used as the between subject factor and copulation duration [difference between copulation duration one and copulation duration five] and total number of laid eggs [difference of total number of eggs female one and total number of laid eggs of female five] were used as covariates (differences had to be taken because SPSS 11.0 could not handle the separated covariates for each repeated measurement). Treatment had no significant effect on the number of

hatched eggs ($F_{1,2}$ = 0.030; P = 0.440; one-tailed test). Treatment x first or fifth copulation interaction had no significant effect on the number of hatched eggs ($F_{1,2}$ = 0.754; P = 0.238; one-tailed test, see also Figure 7). Both covariates were not significant (copulation duration $F_{1,2}$ = 0.085; P = 0.798; number of laid eggs $F_{1,2}$ = 0.019; P = 0.903). When the non-significant covariates were removed, the conclusions did not change.

Because of problems associated with the two fixed covariates, an additional ANCOVA analysis was conducted. Here I used the covariates for each copulation separately, but first and fifth copulation were not treated as repeated measurements but as independent events. Treatment and first or fifth copulation were used as the factors and the number of hatched eggs [arcsine square root transformed proportion of hatched eggs] as the dependent. I included copulation duration and the total number of laid eggs as covariates. There was a close to significant treatment x first or fifth copulation effect on the number of hatched eggs ($F_{1,6}$ = 3.681; P = 0.052; one-tailed test). Copulation duration had a significant effect as covariate ($F_{1,6}$ = 6.003; P = 0.050; longer copulation lead to a higher proportion of fertilised eggs) but the total number of laid eggs was not significant as covariate ($F_{1,6}$ = 1.476; P = 0.270).

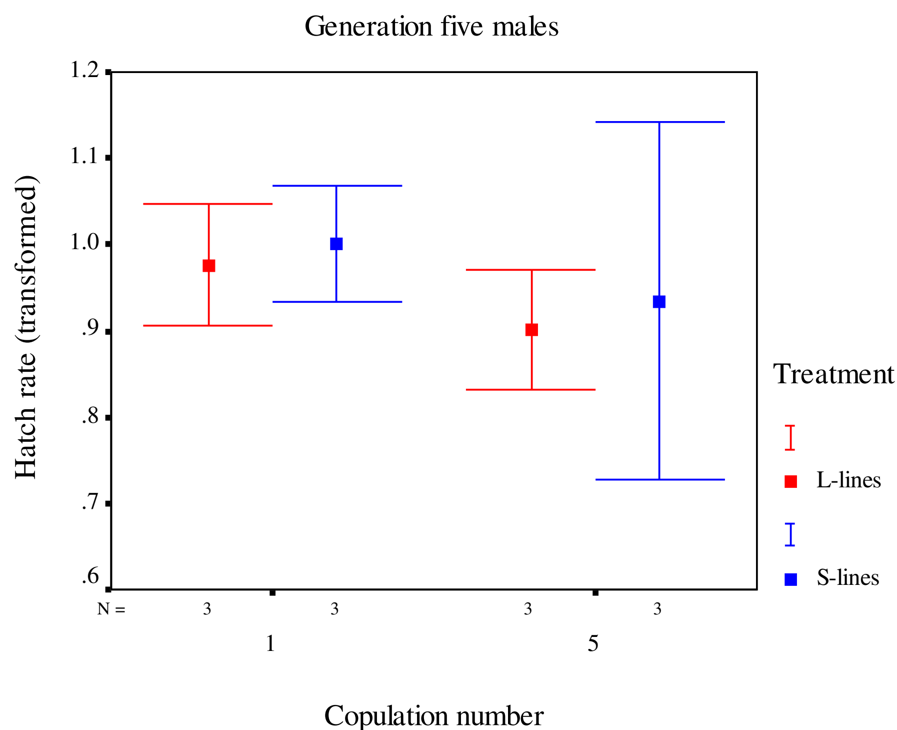
Last I compared the number of hatched eggs [arcsine square root transformed proportion of hatched eggs] for both treatments in the first and in the fifth copulation separated. I did an ANCOVA with number of hatched eggs as the dependent, treatment as the factor and copulation duration and total number of laid eggs were the covariates. Data was split by first or fifth copulation for the analysis. In the first copulation there was no effect of treatment on number of hatched eggs ($F_{1,2}$ = 1.052; P = 0.207; one-tailed test). Both covariates were not significant (copulation duration $F_{1,2}$ = 1.107; P = 0.403; total number of laid eggs $F_{1,2}$ = 0.518; P = 0.547). For the fifth copulation results were the same. Treatment had no significant effect on the number of hatched eggs ($F_{1,2}$ = 0.125; P = 0.378; one-tailed test). Neither copulation duration ($F_{1,2}$ = 4.883; P = 0.158) nor the total number of laid eggs ($F_{1,2}$ = 0.234; P = 0.676) were significant as covariates. Removing the non-significant covariates did not change the conclusion for neither first nor fifth copulation analysis.

Even though one of the three analyses showed a very close to significant treatment x first or fifth

copulation interaction, all together these analyses indicated there was no significant effect of selection on sperm length on fertility. There was no evidence that S-line males had a sperm number advantage in comparison to the L-line males. The hatch rate from copulation one to copulation five changed in the same way for both treatments as there was no significant treatment x first or fifth copulation interaction (Figure 7).

In addition, in the S-lines 5 ± 0.577 males (n=10) had five completed copulations, in the L-lines the mean was 7 ± 1.000 completed fifth copulations (n=10). The number of complete copulations was not significant different between the two treatments, although there was a weak trend for L-line males to have more complete fifth copulations (ANOVA with number of completed fifth copulation as the dependent and treatment as the factor; $F_{1,4} = 3.000$; $P = 0.158$). This also indicated that S-line males and L-line males were of similar quality concerning the size of their sperm reserves. However, when males did not complete five copulations, S-line males stopped earlier than L-line males (ANOVA with number of copulations when not five completed as the dependent and treatment as the factor; $F_{1,4} = 8.939$; $P = 0.040$; S-line males 1.733 ± 0.392 (SE) copulations, L-line males 3.333 ± 0363 (SE) copulations). Together these results suggest more variation in the quality of S-line males as they had a reduced ability to complete more copulations.

Spermatheca size did not differ between the treatments (see below, correlated responses in female reproduction organs). This indicated that L-line females may store less sperm than S-line females due to the sperm length difference and the equal volume of the spermatheca. So the lack of difference between the treatments in fertility above was possibly due to female evolution in the selection lines. L-line females may be more efficient in sperm use. To assess these males from both treatments were mated with wild females which could not co-evolve in the lab. Number of hatched eggs [arcsine square root transformed proportion of hatched eggs] was used as the dependent in an ANCOVA. Treatment was the factor and copulation duration and total number of laid eggs were the covariates. There was no significant effect of treatment on the number of hatched eggs ($F_{1,2} = 8.393$; $P = 0.051$; one-tailed test), but a strong trend for S-line males had a higher proportion of fertilised eggs. Both covariates were not significant (copulation duration $F_{1,2} = 17.214$; $P = 0.053$; total number of laid eggs $F_{1,2} = 4.842$; $P = 0.159$). When the non-significant covariates were removed **Figure 7:**



When mated with females from the same treatment (blue S-lines, red L-lines) both kind of males have the same hatch rate. The mean (± SE) hatch rate did not differ between the two treatments.

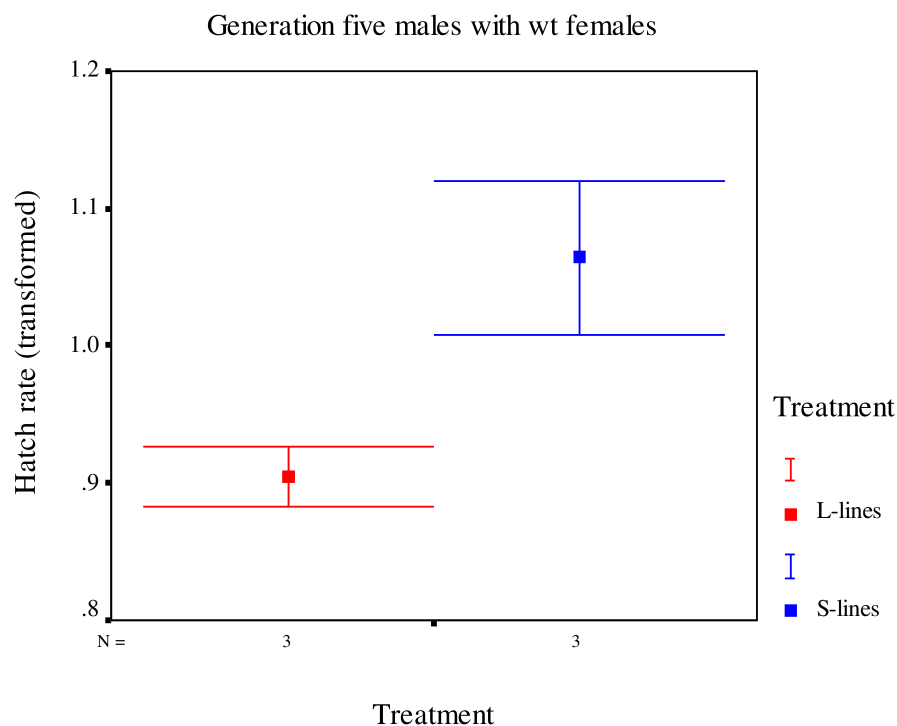


Figure 8: Mean (± SE) hatch rate of both treatment males (blue S-line, red L-line) was significantly different when they were mated with wild females (P = 0.028).

from the model, the trend mentioned just above became significant ($F_{1,4}$ = 7.135; P = 0.028; one-tailed test). This is strong evidence that L-line males have a cost in fertility when the females they mate did not co-evolve with them.

Overall results of the fertility experiments indicated that short and long sperm males had the same fertilisation success when mated with females from the same treatment (Figure 7). When mated with wild females S-line male had a better fertilisation success than L-line males (Figure 8).

Immune System

ANCOVA was used to assess the effect of treatment (factor) on the immune system (dependent). The activity level of phenyleoxidase (PO) [V_{max}] was used as a measure of immune system function. HTL and age were used as the covariates. There was no significant effect of treatment on the immune system ($F_{1,2}$ = 0.177; P = 0.715) and both covariates were not significant either (HTL $F_{1,2}$ = 0.277; P = 0.651; age $F_{1,2}$ = 0.079; P = 0.805). The conclusions remained unchanged when the non-significant covariates were removed. S-line males had a mean PO activity of 382.058 ± 66.007, L-line males had a mean PO activity of 450.614 ± 91.847. It therefore seems, that individuals with longer sperm had no reduced immune system activity.

4.3 Correlated responses in males

As additional potential correlated responses to selection, I looked at some male traits. Those were body size (HTL) and testis volume.

Body size

To see if body size (HTL) evolved in response to sperm length selection I used ANOVA with treatment as the factor and HTL as the dependent variable at generation four. Results indicated there was a weak trend for L-line males to become larger as a result of selection on sperm size ($F_{1,4}$ = 8.829x10^{-3}; P = 0.186). S-line males had a mean HTL of 3.478 ± 0.010 mm and L-line males had a HTL mean of 3.524 ± 0.026 mm (Figure 9).

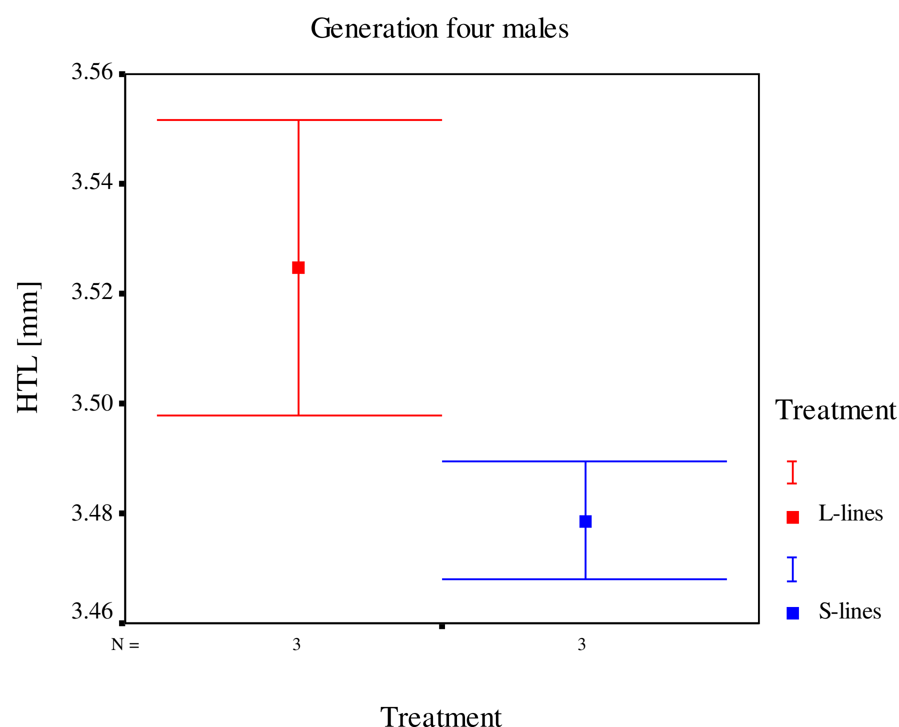
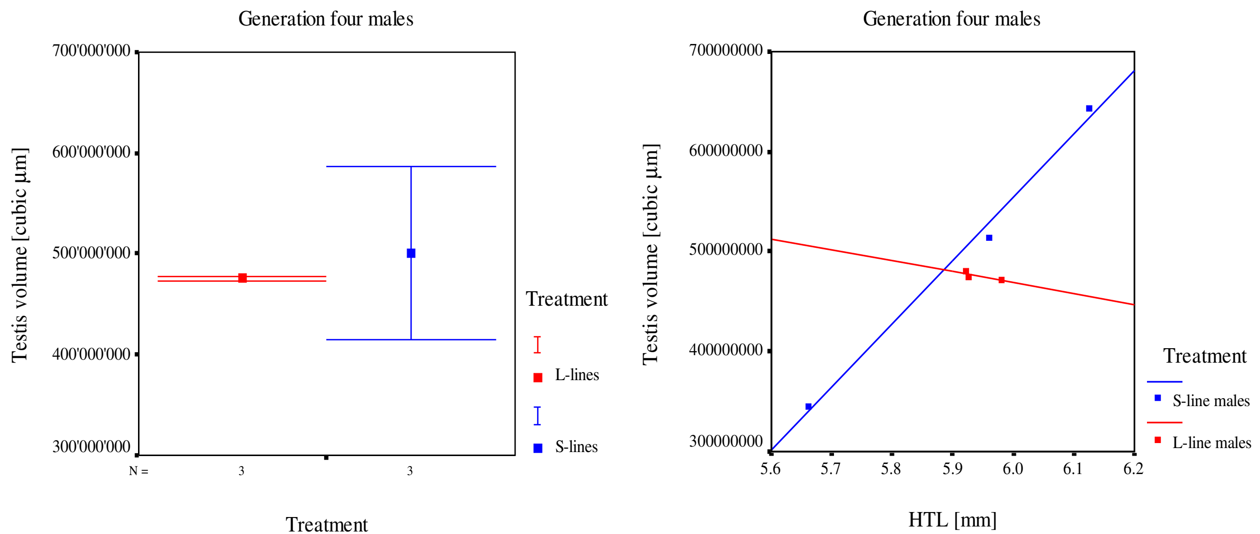


Figure 9: Mean HTL (± SE) [mm] for males of both treatments (blue S-lines, red L-lines) at generation four. Treatments did not differ significantly (P = 0.186).

Testis volume

To investigate the effect of selection on testis volume when males were fully mature (14 to 20 days after emergence), I used a ANCOVA with treatment as the factor and testis volume [(testis volume)$^{0.33}$ µm] as the dependent, and I used HTL as the covariate. Unfortunately age was not recorded for each single fly in this generation (generation four). Therefore I could not use age as a



covariate in **Figure 10:** Left: Testis volume (± SE) [µm^3] was not different in the two treatments (blue S-lines, red L-lines) at generation four. Right: S-line males at generation four showed a positive testis volume / body size correlation whereas L-line males had all about the same testis volume independent of their body size.

the analysis. Treatment had no significant effect on testis size ($F_{1,3}$ = 433.461; P = 0.159). Testis volume mean for S-lines was 500.785x10^6 ± 2.603x10^6 µm^3 and 475.351x10^6 ± 86.217x10^6 µm^3 for L-lines. The covariate HTL was significant associated with testis volume, as larger males had larger testis ($F_{1,3}$ = 103.629; P = 0.002) (Figure 10). However there was a significant treatment x HTL interaction effect ($F_{1,2}$ = 351.221; P = 0.031; see also Figure 10), indicating that S-line males had a positive body size / testis size correlation, whereas L-line males had a (constant) testis volume independent of the body size.

4.4 Correlated responses in females

As potential correlated responses to selection in females, I looked at female body size (HTL), developmental time, egg size, size of reproduction organs and the PO activity level (as an important part of the immune system). The reproduction organs I looked at were the spermatheca, the spermathecal duct and the accessory glands.

Body size

ANOVA with treatment as the factor and HTL as the dependent showed a significant effect of treatment on HTL ($F_{1,4} = 8.954$; $P = 0.040$) in generation four. S-line females were smaller than L-line females. S-line females had a HTL mean of 2.837 ± 0.008 mm and L-line females had a HTL mean of 2.903 ± 0.020 mm (Figure 11). Results indicated that selection caused a significant correlated response in female body size (but not in males).

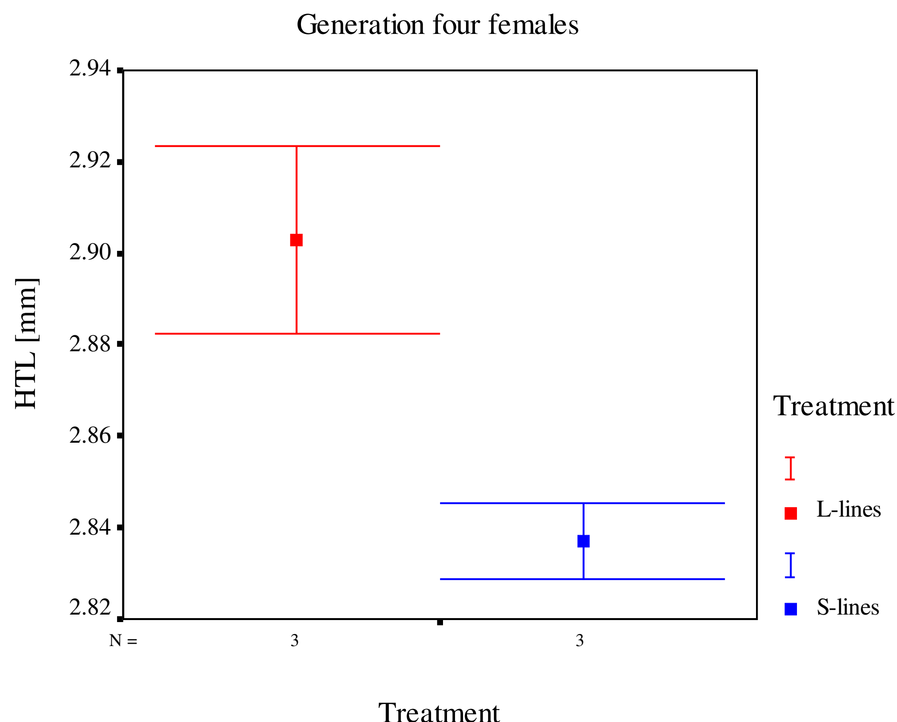


Figure 11: Mean HTL (± SE) [mm] for females of both treatments (blue S-lines, red L-lines) at generation four. Other than in males had selection a significant effect on body size (P = 0.040).

Developmental time

Developmental time was examined in females in the same way as in the males. ANOVA was used to assess effects of bidirectional selection on developmental time in females. Developmental time in generation four was used as the dependent and treatment as the factor. HTL was used as the covariate. Result showed no significant effect of treatment on developmental time in generation four ($F_{1,3}$ = 0.755; P = 0.225; one-tailed test). HTL was not significant as covariate ($F_{1,3}$ = 2.439; P = 0.216). S-line females had a mean developmental time of 24.534 ± 0.416 days, L-line female had a mean of 25.690 ± 0.391 days (Figure 12). When the non-significant covariate was removed, the result indicated S-line females tended to have a shorter developmental time than L-line females ($F_{1,4}$ = 4.086; P = 0.057; one-tailed test).

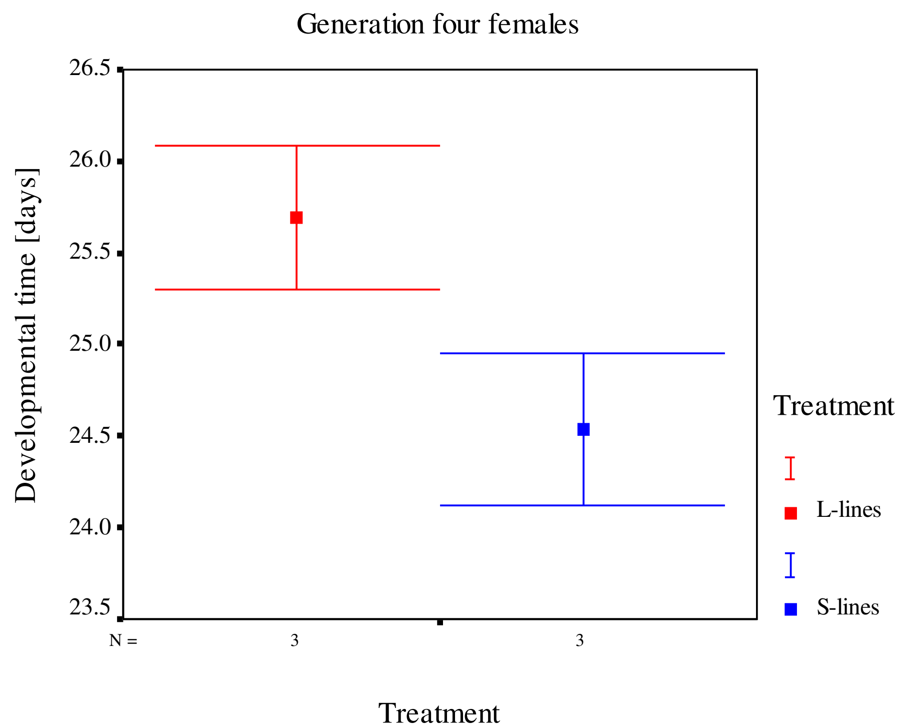


Figure 12: Developmental time (± SE) [days] for the females of both treatments (blue S-lines, red L-lines) at generation four. Developmental time between the treatments was not significantly different (P = 0.057).

To further assess whether there may nevertheless be a difference in the developmental time of the two treatments an additional analysis was employed to further test these data. This was done in the same way as in the males. I used linear regression techniques and compared the slopes of the two

linear regressions (split by lines: i.e. one for S-lines and one for L-lines) of developmental time (Y-axis) against generation (X-axis), as these make use of the data from three rather than a single generation. Slopes were compared using the methods of Zar (1998). Again while the slopes of these regression were in the predicted direction ($slope_S = 0.024$; $slope_L = 0.358$) and the difference was statistically significant ($|t| = 5.019 > t_{0.05(1),14}$; $P < 0.001$; H_0: $slope_S = slope_L$ had to be rejected) between the two regressions. Although ANOVA showed no significant difference in developmental time in generation four, the slopes of the two regression lines were different. By having a look at Figure 13 of developmental time over generations it appears, that the two treatments would significantly differ in developmental time with several more generations of selection assuming the current trend is maintained. Females of the S-lines tend to have a shorter developmental time than females from the L-lines. This was the same result as found in the males. The slopes of the linear regression were significantly different in both sexes, indicating that the effect on developmental time exists in both sexes.

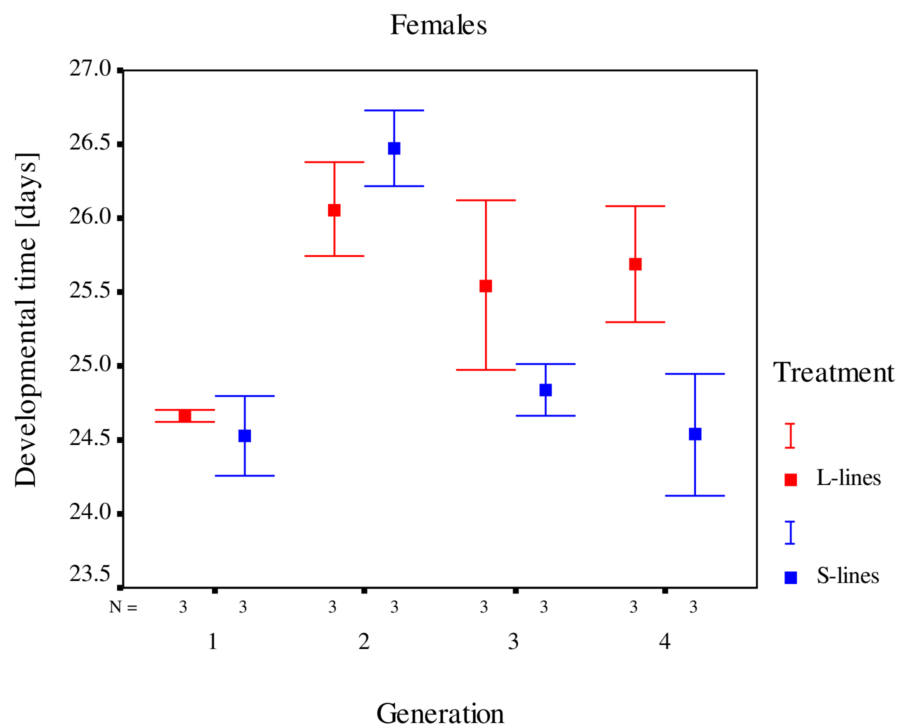


Figure 13: Mean (± SE) developmental time [days] for females of both treatments (blue S-lines, red L-lines) over the four generation with selection. Generation two was excluded from the analyses because the rearing temperature was only 13° Celsius whereas all other generation were reared at 20° Celsius.

Egg size

To assess the correlated response on female gamete size due to selection on male gamete size I used an ANCOVA with egg volume [(egg volume)$^{0.33}$ µm] as the dependent variable, treatment as the factor and HTL as the covariate. There was no significant effect of treatment on egg volume ($F_{1,3} = 0.006$; $P = 0.943$; mean egg volume ± SE: S-lines 72.135x10^{6} ± 0.739x10^{6} µm^{3}; L-lines 73.448x10^{6} ± 1.329x10^{6} µm^{3}; see also Figure 14). HTL was not significant as covariate, but there was a trend for larger females to have larger eggs ($F_{1,3} = 8.155$; $P = 0.065$). When the non-significant covariate was removed the conclusions stayed the same. Selection on male gamete length had no significant effect on female gamete size.

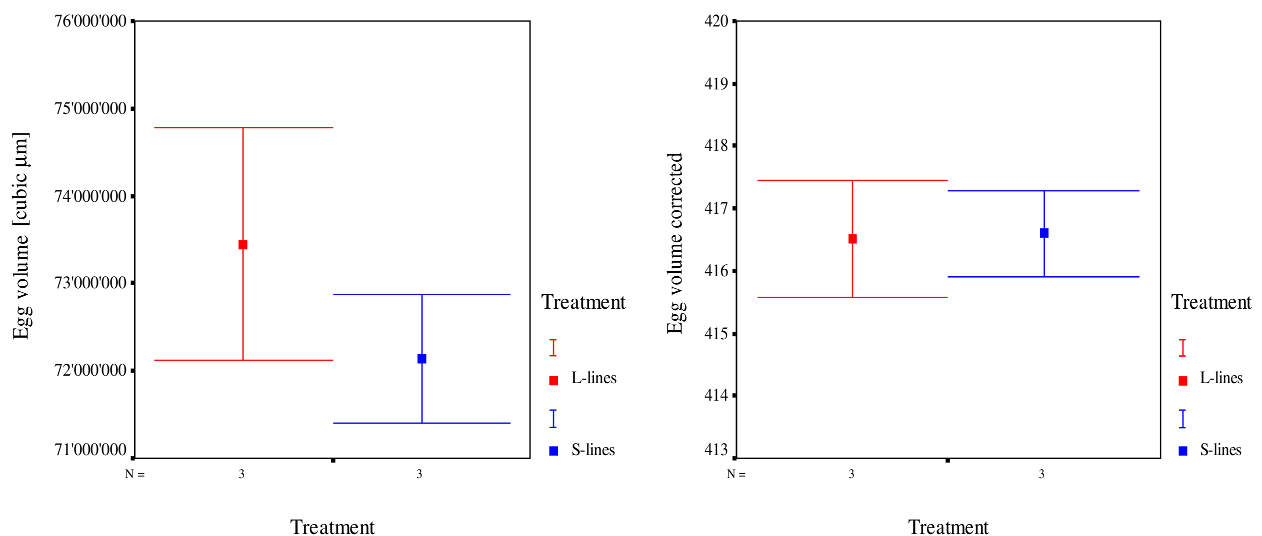


Figure 14: Left: Egg volume (± SE) [µm^{3}] for females of both treatments (blue S-lines, red L-lines) at generation five. Right: Egg volume (± SE) corrected for HTL.

Reproduction organs

ANCOVA was then used to examine the effect of treatment on correlated responses in female reproduction organs morphology trait separately. I first used spermatheca area [(spermatheca area)$^{0.5}$ µm] as the dependent variable, treatment as the factor and HTL as the covariate. There was no significant effect of treatment on spermatheca area ($F_{1,3} = 2.224$; $P = 0.233$; mean area ± SE: S-lines 12.452x10^{3} ± 0.303x10^{3} µm^{2}; L-lines 12.360x10^{3} ± 0.625x10^{3} µm^{2}). HTL was not significant as covariate ($F_{1,3} = 2.825$; $P = 0.191$). Removal of the non-significant covariate did not change the

conclusion.

Secondly, I assessed effects of bidirectional selection on spermathecal duct length. ANCOVA was used with spermathecal duct length [µm] as the dependent, treatment as the factor and HTL as the covariate. There was no significant effect of treatment on spermathecal duct length ($F_{1,3} = 0.051$; P = 0.418; one-tailed test; mean length ± SE: S-lines 747.448 ± 12.429 µm; L-lines 770.237 ± 16.075 µ m). Although this results showed no sign of response, a look on the graph (Figure 15) points out a

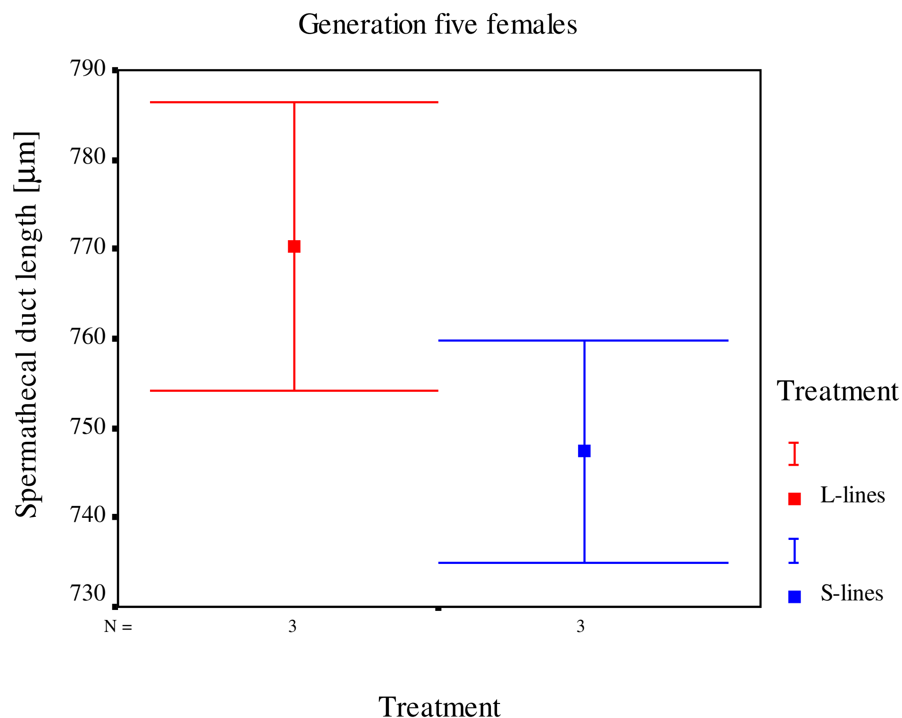


Figure 15: Mean spermathecal duct length (± SE) [µm] for females of both treatments (blue S-lines, red L-lines) at generation five.

correlated response to selection seemed to go in the predicted direction. The females from the L-lines tended to have longer ducts than the females from the S-lines. This might be caused by the possible necessary correlation between sperm length and female duct length (see Minder *et al.* in press). Significant difference might be reached with further generation of selection. HTL was not significant as covariate ($F_{1,3} = 0.156$; P = 0.719). When the non-significant covariate was removed, the conclusion did not change. But the results indicated some kind of trend for a difference between the two treatments ($F_{1,4} = 1.258$; P = 0.167; one-tailed test). S-line females tended to have shorter spermatheca ducts than L-line females (see also Figure 15).

Finally, I examined the accessory gland. Again ANCOVA was used to assess correlated response. I used accessory gland area [(accessory gland area)$^{0.5}$µm] as the dependent, treatment as the factor and HTL as the covariate. There was no significant effect of treatment on accessory gland area ($F_{1,3}$ = 3.826; P = 0.145; mean area ± SE: S-lines 105.602x10^{3} ± 3.365x10^{3} µm^{2}; L-lines 110.563x10^{3} ± 5.515x10^{3} µm^{2}). In this case HTL as covariate was significant ($F_{1,3}$ = 13.041; P = 0.036; larger females had larger accessory glands) (Figure 16).

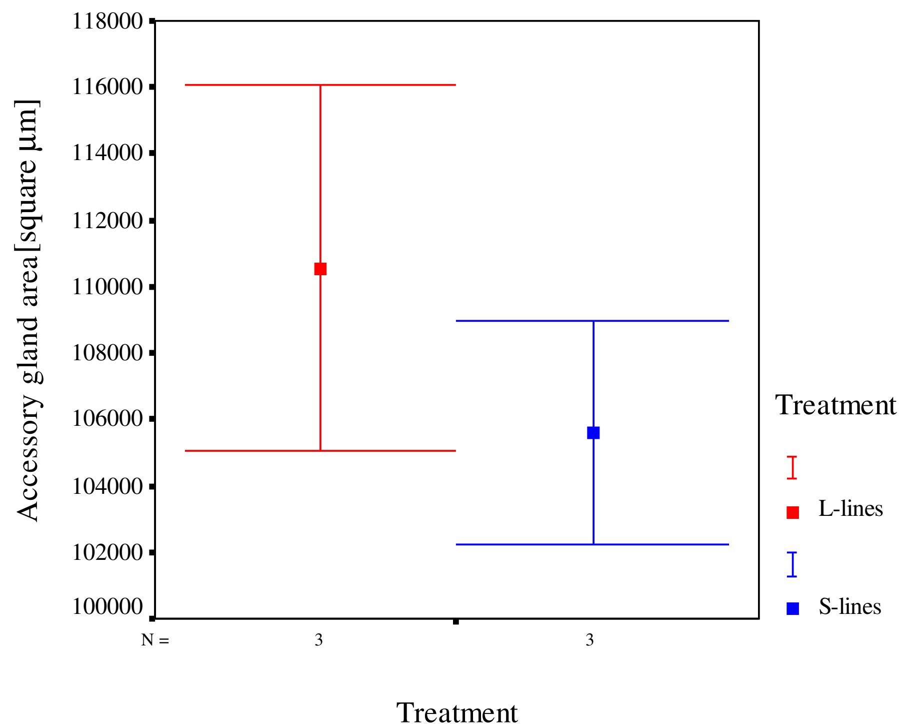


Figure 16: Mean accessory gland area (± SE) [µm^{2}] for females of both treatments (blue S-lines, red L-lines) at generation five.

Although results showed no significant correlated responses in females, results indicated that with more generations of selection responses would become significant. At least in the spermatheca duct length a trend in the predicted direction was observable with L-line females tended to get longer spermatheca ducts. Additionally L-line females tended to become larger accessory glands than S-line females. Possible reasons for this have to be discussed.

Immune system

ANCOVA was used to assess the correlated response of the immune system (PO activity) to the two different sperm selection treatments. PO activity was used as the dependent, treatment as the factor. HTL and age were used as the covariates. Results showed no significant effect of treatment on the PO activity ($F_{1,2} = 0.055$; $P = 0.836$). Neither HTL ($F_{1,2} = 2.917$; $P = 0.230$) nor age ($F_{1,2} = 2.494$; $P = 0.255$) were significant as covariate. As in the males there was no detectable correlated response. S-line females had a mean PO activity of 540.148 ± 252.027, L-line females had a mean PO activity of 837.951 ± 132.506. When the non-significant covariates were removed, the conclusions remained unchanged.

5. Discussion

Significant difference in sperm length was reached after only four generation of bidirectional selection. This result is consistent with the high heritability of sperm length in *S. stercoraria* found by Ward (2000). As responses to selection are commonly asymmetrical in quantitive traits, the unequal response for short sperm selection and long sperm selection was no big surprise. But the kind of the asymmetry was a little bit surprising. Longer sperm were assumed to have costs for the males, so why should the response to selection be better than in the S-lines? In fact the observed asymmetry may not be an effect due to the selection *per se*. It is possible that the asymmetry I found was an artificial effect of lab rearing. It is known, that *S. stercoraria* respond with longer sperm to higher temperature (Hellriegel & Blanckenhorn 2002). The flies came from the field in mid October, where temperature was relatively low compared to the lab temperature. Moreover sperm from both treatments became longer in the first selection generations. Even at generation four S-line males had still longer sperm than the males caught in the field. So temperature was probably the reason for this asymmetric response.

Several potential costs for longer sperm and correlated responses to selection on sperm length were investigated. Although I selected on sperm length for only four generations, some strong trends and significant results were found. While developmental time was not significantly different between the two treatments at generation four, looking at the developmental time over the generations showed development time had significantly diverged in the predicted direction for males and females. L-line developmental time increased much more than S-line developmental time. Essentially the same results were found by Pitnick *et al*. (1995) for males in *Drosophila spp*. As I found no difference in the longevity for both treatments, the longer developmental time is a cost for the males (and females) of the L-lines. A shorter developmental time is though to be advantageous in *S. stercoraria* because dung is limited at this stage due to strong intra- and interspecific food competition, predation and strong seasonal effects (larvae have to complete the development to over-winter or because cow pats may dry out) (Hosken *et al*. 2003 and references therein). And a shorter developmental time is additionally thought to be advantageous because *S. stercoraria* have multiple generations per year (Hosken *et al.* 2003).

There was no sign for a sperm length / sperm number trade-off as the fertilisation rate (measured as the proportion of hatched eggs) was not different between the two treatments when males and females from the same lines were mated. Females of *S. stercoraria* receive enough sperm to fertilise four full clutches with one full copulation (Parker 1970). In my experiment females laid only three to four clutches on average before they died (data not shown). But looking only on females with five (or more) clutches did not change the results. L-line males even tended to have more completed fifth copulations than S-line males. However it should be noted, that the females in this fertility assessment had the possibility to adapt to selection on sperm size. If sperm length and sperm number did trade-off, L-line females may have been indirectly selected to use sperm more efficiency. In addition, wild females mated to S-line males had a higher egg hatch rate than those mated to L-line males. However sperm length / female reproduction organs associations can influence fertility (Gage & Morrow 2003, Miller & Pitnick 2002). For example, in *D. melanogaster* longer sperm have a higher fertilisation rate when the female seminal receptacle (sperm storage organ) is larger (Miller & Pitnick 2002). In *S. stercoraria* sperm length /female reproduction organ size (spermathecal duct) correlation may also to influence fertilisation rate. Because S-line males had sperm of about the length of wild males and only the L-line males had really different sperm from wild flies. However the decreased fertilisation rate of L-line males mated with wild females is clear, although its interpretation is unclear based on all the above. One possibility is, that L-line females have increased sperm use efficiency, the other is, that wild females select against longer sperm. This requires more investigation. Additionally, as my measure of sperm number was only indirect, it is still possible, that males from both treatments produced different numbers of sperm. In retrospect, counting the number of sperm bundles might have been a better way to assess differences in sperm number between the two treatments (see Pitnick 1996).

Another surprising result was, that L-line females were significantly larger than S-line females. A weak trend for the same was also found in males. This finding contrasts somewhat with other data on body size / sperm length relationship. Sperm length is independent of body size in *S. stercoraria* (e.g. Ward 1998, Ward & Hauschteck-Jungen 1993) and hence any kind of explanation for this unexpected finding is purely speculative. If selection for longer sperm also selected for larger reproduction organs in L-line flies (weak trend in spermathecal duct length and accessory gland

size), it is possible that selection inadvertently increased body size as well. This would then also explain the rapid and steep increase in developmental time of the larger L-line flies.

Results also showed a interesting treatment x body size interaction for testis volume. In S-line males testis volume was positively correlated to body size (as expected, see Hellriegel & Blanckenhorn 2002), whereas in L-line males testis volume appears to be largely independent of body size. This result may be partly driven by the small variation in body size means in the L-lines and does not seem to hold within lines (data not shown). However it seems to be possible that selection on sperm length changed something about the testis size allometry. A possible reason is that L-line males need larger testis to avoid a sperm length / sperm number trade-off (see Pitnick & Markow 1994a). Therefore they may initially have to have testis of a certain minimal size independent of body size.

In any case there was no longevity costs for males with longer sperm nor was the PO activity level different in males or females from the two treatments. This is surprising as one of the expected key trade-offs in evolution would be between reproduction and longevity (e.g. Stearns 1992). Males from both treatments reached sexual maturity at the same age post emergence, although testis growth rate was higher in the L-lines, even though at individual ages (except day1 post emergence when testis size of L-males was greater) there was considerable overlap in testis size between the treatments. Testes volume was also the same for older males of both treatments (14 to 20 days old). Pitnick *et al*. (1995) found, that males with longer sperm had a delay in maturity in *Drosophila spp*. In *Drosophila spp*. the testes length is about the length of the sperm (Lindsley & Tokuyasu 1980) but in *S. stercoraria* sperm are many time smaller than the testis. To reach a minimal testis length depending on the sperm length is therefore not thought to be the crucial factor for the achievement of sexual maturity in *S. stercoraria* (see also Foster 1967), other than in *Drosophila spp*. (Pitnick *et al*. 1995). This is consistent with the observation that both treatments had the same final testis volume (in spite of growth rate differences across days) and reached maturity at the same age. The higher testis growth rate of the L-line males has some possibilities: L-line males may eat more, they may be more efficient in energy conversion or they have some trade-off somewhere else. In summary, there is some evidence that males may pay some costs in term of testes growth rate and age at maturity in *S. stercoraria* when they have longer sperm. The increased developmental time

(discussed above) possible due to the larger testis at the first day after emergence may not be the only cost.

Egg size did also not differ between the two treatments. The most likely explanation is that sperm size and egg size are NOT genetically correlated at all. As far as I know is there only one study that found a correlation between sperm length and egg size and that was for frogs (Byrne *et al*. 2003). However tend larger females to have larger eggs. Like in the males there is a positive gamete size / body size correlation in females. In the context of this selection experiment, the reason why is not clear, as in insects egg number (rather than egg size) is generally positive associated with body size (e.g. Parker 1970, Minder *et al.* in press). Furthermore, even if the sperm length and egg size were related, if egg size changes were of the same magnitude as sperm length changes, a difference may not be detectable. Also none of the investigated female reproduction organs showed any significant correlated response to the selection on sperm length. Spermatheca showed no correlated response although Otronen *et al*. (1997) found, that longer sperm are more likely to be stored in the spermatheca of female *S. stercoraria*. So L-line females should not be able to store as many sperm as S-line females. This would be good evidence for L-line females use sperm more efficiency and would explain why I found no sperm length / sperm number trade-off. Miller and Pitnick (2000) found a positive correlation between the sperm length and the seminal receptacle in *D. melanogaster*. Spermathecal duct showed at least a weak trend for correlated response in *S. stercoraria* as L-line females tended to have longer spermathecal ducts. A positive sperm length / spermathecal duct length correlation is already known for *Scathophagidae* (Minder *et al*. in press). There was also the same kind of trend for the accessory glands. But the exact meaning of this result remains unclear. The exact function of the accessory glands has not been established. It is thought, that these glands release a fluid during copulation to lubricate (Hosken *et al*. 2001). Or may be the glands receive substances from the male during copulation (Hosken & Ward 1999). Unfortunately I only measured females after copulation but no virgin females. So no clear conclusion can be made about the accessory glands. But for all female reproduction organs may obtain the same as mentioned before. Selection for longer sperm possibly also selects for larger reproduction organs due to a correlated selection on body size.

To summarise I found increased developmental time and reduced fertility when mated with wild females as cost for males with longer sperm. Females from L-lines showed the same increase in developmental time. They also had a weak tendency to have longer spermathecal ducts as a correlated response to selection. S-line males seem to have a body size-dependent testis volume whereas L-line males did not. This needs further examinations. As selection took place for only four generations, the achieved divergence may not be great enough for other costs to be manifest. Nevertheless like in *Drosophila spp.* there seem to be some costs related to production of longer sperm in the yellow dung fly *S. stercoraria*.

6. Acknowledgements

First I would thank Dr. David J. "Master D" Hosken for the excellent supervision. I'd like to thank him for the help in the lab and the music he brought there... and for his fast response to all the previous versions of this thesis I sent him via e-mail. And of course for the motivation all over the last year.
I would also thank Prof. Dr. Paul I. Ward that gave me the opportunity to make my diploma thesis in the Zoological Museum.
Gioia Schwarzenbach gave logistic support over the whole time of the experimental phase of my thesis and useful hints over the whole time.
Dr. Wolf U. Blanckenhorn gave supporting help in the weird field of statistics...
Rahel Leugger a big "thank you" for sharing the same fate as me.
Eline Embrechts, Philip Schönenberger, G.S. and R.L. were good helping hands in the hard weeks of the experiments.
Marco Demont, Thomas Bucher and many others had always time for a coffee break. As you all might know: life without coffee breaks would be unbearable!
It was also nice to share the diploma room with Doina Muncaciu, Manuela Varini, M.D. and R.L. Think we had a lot of fun...
At this place also thank to my parents, Leo and Maria Dobler who always supported my plans to study, even when the success was not on the hand. And to my brother Andreas, it's nice to have a little brother like you!
Think I would never have come this far without the learning group (Cornelia Haefeli, Christoph Meier, Susanne Keller and Samuel Wüst) for the exams after the second year. I thank and hug you C.H, C.M., S.K and S.W.
It was not easy to find a suitable title for my thesis talk. But with Sandra Hangartner, R.L., Facitas *ad libitum* and three bottles of red wine it was found: Super Size Me?! - The Costs of Longer Sperm in the Yellow Dung Fly *Scathophaga stercoraria*.
Not to forget "Therapy?" for their 1994 album "Troublegum", was always better than a cup of black coffee at 9 pm in the lab.

7. References

Amano K (1983); Studies on the intraspecific competition dungbreeding flies. I. Effects of larval density on the yellow dung fly, *Scathophaga stercoraria* L. (Diptera: Scathophagidae); Japanese Journal of Sanitary Biology; **34**:165-175

Baccetti B and Dallai R (1978); The spermatozoon of arthropoda XXX. The multiflagellate spermatozoon in the termite *Mastotermes darwiniensis*; The Journal of Cell Biology; **76**:569-579

Blanckenhorn WU (1997); Effects of temperature on growth, development and diapause in the yellow dung fly – against all the rules?; Oecologia; **111**:318-324

Blanckenhorn WU and Hellriegel B (2002); Against Bergmann's rule: Fly sperm size increases with temperature; Ecology Letters; **5**:7-10

Byrne PG, Simmons LW and Dale Roberts J (2003); Sperm competition and the evolution of gamete morphology in frogs; Proceedings of the Royal Society London B; **270**:2079-2086

Falconer DS and Mackay TFC (1996); Introduction to quantitative genetics, 4th Edition; ISBN 0582243025; Longman; pp 184-207

Foster W (1967); Hormone-mediated nutritional control of sexual behaviour in male dung flies; Science; **158**:1596-1597

Gage MJG (1998); Mammalian sperm morphometry; Proceedings of the Royal Society London B; **265**:97-103

Gage MJG and Morrow EH (2003); Experimental evidence for the evolution of numerous, tiny sperm in sperm competition; Current Biology; **13**:754-757

Gomendio M and Roldan ERS (1991); Sperm competition influences sperm size in mammals; Proceedings of the Royal Society London B; **243**:181-185

Hellriegel B and Blanckenhorn WU (2002); Environmental influences on the gametic investment of yellow dung fly mates; Evolutionary Ecology; **16**:505-522

Hosken DJ (2001); Sex and death: Microevolutionary trade-offs between reproductive and immune investment in dung flies; Current Biology; **11**:R379-R380

Hosken DJ (2003); Sperm Biology: Size indeed matters; Current Biology; **13**:R355-R356

Hosken DJ, Garner TWJ, Tregenza T, Wedell N and Ward PI (2003); Superior sperm competitors sire higher-quality young; Proceedings of the Royal Society London B; **207**:1933-1938

Hosken DJ, Garner TWJ and Ward PI (2001); Sexual conflict selects for male and female reproductive characters; Current Biology; **11**:489-493

Hosken DJ, Uhía E and Ward PI (2002); The function of female accessory reproductive gland secretion and a cost to polyandry in the yellow dung fly; Physiological Entomology; **27**:87-91

Hosken DJ and Ward PI (1999); Female accessory reproductive gland activity in the yellow dung fly *Scathophaga stercoraria* (L.); Journal of Insect Phyiology; **45**:809-814

Lindsley DL and Tokuyasu KT (1980); Spermatogenesis; The genetics and biology of *Drosophila* Volume 2 (eds. Ashburner M and Wright TRF); ISBN 0120649438; Academic Press; pp 225-294

Miller GT and Pitnick S (2002); Sperm-female coevolution in *Drosophila*; Science; **298**:1230-1233

Minder AM, Hosken DJ and Ward PI (in press); Co-evolution of male and female reproductive characters across the Scathophagidae (Diptera); Journal of Evolutionary Biology

Morrow EH and Gage MJG (2001); Sperm competition experiments between lines of crickets producing different sperm lengths; Proceedings of the Royal Society London B; **268**:2281-2286

Oppliger A, Hosken DJ and Ribi G (1998); Snail sperm production characteristics vary with sperm competition risk; Proceedings of the Royal Society London B; **265**:1527-1534

Otronen M, Piedad R and Ward PI (1997); Sperm storage in the yellow dung fly *Scathophaga stercoraria*: Identifying the sperm of competing males in separate female spermatheca; Ethology; **103**:844-854

Parker GA (1970); Sperm competition and its evolutionary effect on copulation duration in the fly *Scatophaga stercoraria*; Journal of Insect Physiology; **16**:1301-1328

Parker GA (1982); Why are there so many tiny sperm? Sperm competition and the maintenance of two sexes; Journal of Theoretical Biology; **96**:281-294

Parker GA (1993); Sperm competition games: Sperm size and sperm number under adult control; Proceedings of the Royal Society London B; **253**:246-254

Parker GA and Simmons LW (1994); Evolution of phenotypic optima and copula duration in dung flies; Nature; **370**:53-56

Pathak JPN (1993) Insect immunity; ISBN 0792320867 ;Kluwer Academic Press

Pitnick S (1996); Investment in testes and the cost of making long sperm in *Drosophila*; The American Naturalist; **148**:57-80

Pitnick S and Markow TA (1994a); Male gametic strategies: Sperm size, testes size, and the allocation of ejaculate among successive mates by the sperm-limited fly *Drosophila pachea* and its relatives; The American Naturalist; **143**:785-819

Pitnick S and Markow TA (1994b); Large-male advantages associated with costs of sperm producing in *Drosophila hydei*, a species with giant sperm; Proceedings of the National Academy of Science of the United States of America; **91**; 9277-9281

Pitnick S, Markow TA and Spicer GS (1995); Delay male maturity is a cost of large sperm in *Drosophila*; Proceedings of the National Academy of Sciences of the United States of America; **92**:10614-10618

Pitnick S and Miller GT (2000); Correlated response in reproductive and life history traits to selection on testis length in *Drosophila hydei*; Heredity; **84**:416-426

Schwarzenbach GA, Hosken DJ and Ward PI (in press); Sex and immunity in the yellow dung fly *Scathophaga stercoraria*; Journal of Evolutionary Biology

Sigurjónsdóttir H (1980); Evolutionary aspects of sexual dimorphism in size: Studies on dung flies and three groups of birds; PhD Thesis, University of Liverpool

Simmons LW, Stockley P, Jackson RL and Parker PA (1996); Sperm competition or sperm selection: no evidence for female influence over paternity in yellow dung flies *Scatophaga stercoraria*; Behavioural Ecology and Sociobiology; **38**:199-206

Söderhäll K and Cerenius L (1998); Role of prophenoloxidase-activating system in invertebrate immunity; Current Opinions in Immunology; **10**:23-28

Stearns SC (1992); The evolution of life histories; ISBN 0198577419; Oxford University Press; pp 72-90

Ward PI (1998); Interspecific variation in sperm size characters; Heredity; **80**:655-659

Ward PI (2000); Sperm length is heritable and sex-linked in the yellow dung fly (*Scathophaga stercoraria*); Journal of the Zoological Society of London; **251**:349-353

Ward PI and Hauschteck-Jungen E (1993); Variation in sperm length in the yellow dung fly *Scathophaga stercoraria* (L); Journal of Insect Physiology; **39**:545-547

Ward PI and Simmons LW (1991); Copulation duration and testes size in the yellow dung fly, *Scathophaga stercoraria* (L.): The effects of diet, body size, and mating history; Behavioural Ecology and Sociobiology; **29**:77-85

Yaeger DP and Ultsch GR (1989); Physiological regulation and confirmation: A BASIC program for the determination of critical points; Physiological Zoology; **62**:888-907

Zar JH (1998); Biostatistical analysis, 4th Edition; ISBN 0130823902; Prentice Hall International Editions; pp 360-368