Estimating genetic parameters of larval and pupal development time in *Hydrotaea aenescens* (Diptera: Muscidae)

David Wongso
REG.NR.: 1012535

Major thesis Animal Breeding and Genomics (ABG 80436)
January 2023

Supervisors/examiner:
Aniek Bouwman
Esther Ellen
Piter Bijma

Thesis: Animal Breeding and Genomics

WAGENINGEN
UNIVERSITY & RESEARCH

Abstract

Black dump fly (BDF; *Hydrotaea aenescens*/*Ophyra aenescens*) is used as a biological control agent to reduce house fly populations in animal stables. For mass rearing in factory setup, it can be difficult to have a consistent production schedule due to high variation in BDF's development time. The aim of this report is to estimate the genetic parameters of the BDF's larval and pupal development time, and assess the potential to shorten and synchronize development time through breeding programs. A full-sib-half-sib breeding design was used to collect traits measurement, then animal model and sire model were used to compute the genetic parameters. In total 57 sires and 99 families/dams were reared. They produced 5509 larvae, in which 3424 emerged as adults. The standard deviation (SD: 1.8 days) and heritability (h^2: 0.61) of the larval development time is high, in comparison to that of pupal development time (SD: 0.55 days; h^2: 0.05). In addition, the genetic correlation between larva and pupal development time is considerably low (0.12), showing that larval development time has the potential to be improved through selection, while pupal development time does not. However, the data also showed a trend that shorter larval development time may be correlated with lower adult emergence rate, therefore it is recommended to have selection toward a central mean, instead of directional selection toward shortest larval development time.

Contents

1. Introduction

Black dump fly (BDF; *Hydrotaea aenescens,* formerly *Ophyra aenescens*) is a housefly-like insect that originates from Central and South America. It is a holometabolous (complete metamorphosis) insect, which means that it goes through four stages: egg, larva, pupa and adult. BDF's larvae can be commonly found in rotten meat or manure, hence it is an important insect for forensic entomology (Lefebvre & Pasquerault, 2004; Cortinhas et al., 2016). However, it has also been used as biological control of houseflies in the swine stable, and this is due to two characteristics: 1) its larva has predatory behavior towards other fly's larva of similar size, 2) its adult is less aggravating to the animal and human by staying low in the dark corner and does not fly as much. (Hogsette & Jacobs, 2003). Although it has been typically used in swine stables, it also shows potential usage in cattle and poultry stables (Nolan & Kissam, 1987; Hogsette et al., 2002).

Bestico B.V. is a company in the Netherlands that produces fly related bio-control products. Bestico rears BDF and sells live BDF's pupae as a product to farmers. This product, however, still requires optimization in its production and the quality. One of the challenges that Bestico wants to overcome is the synchronizing of harvesting timing. At around eight to ten days after hatching from eggs, the larva will reach ca. 20mg average weight. The larvae are then removed from the feed medium and placed on the harvest net to separate the pupae and larvae. This process is repeated every day for two to three consecutive days. Individuals that do not turn to pupae after three days are discarded and are considered lost. Currently, BDF has high variation in the duration of its life stages, meaning that the losses are quite substantial. Having the BDF synchronized on their life cycle will not only result in less larvae being wasted, it also allows better prediction on production schedule, so that labor planning can be more efficient.

The high variation of duration in different stages of BDF can be due to many reasons, some examples are: 1) Sex, male BDF tends to have shorter development time than female (Johnson & Venard, 1957). 2) Temperature, the higher the temperature, the faster BDF grows, and less variation within the population (Lefebvre & Pasquerault, 2004). 3) Feed composition, higher protein concentration in feed, results in a longer development time (Hogsette & Washington,1995). However, it is unclear how much variation in these traits is influenced by genetics and can be optimized through the means of a breeding program.

It has been shown that breeding programs in insects is possible, as Facchini et al. (2022) had successfully bred black soldier fly (*Hermetia illucens)* for heavier larvae, but

selection of insect lines has its own unique challenges. Unlike conventional farm animals, insects have a faster life cycle and higher fecundity, which means that there is a short timeframe and large number of candidates to handle. The fragility and the small size of insects also make tagging and phenotyping (e.g. tags, notching) difficult. Tracking individual insects for pedigree or consecutive measurements, is usually done by individual rearing, which adds more workload pressure to the already short handling timeframe. These constraints can result in a costly breeding program, therefore knowing the genetic parameters of the traits is important. Genetic parameters show whether a trait has the potential to be selected, and can be used to simulate the response to assess whether a breeding design is worth investing in.

The aim of this study is to explore the possibilities of shortening and synchronizing the BDF's larval and pupal development time by genetic selection, through analyzing its genetic parameters.

2. Materials and methods

2.1 Data collection

Prior to the start of the data collecting, preliminary tests were done on the practicality of the lab rearing. It was found that the optimal mating ratio is one male to two females. More females will lead to a diminishing rate of mating success. The possibility to rear multiple females with a single male indicates that a full-sib-half-sib breeding design can be used. Offspring of the same sire and dam are full-siblings, whereas half-siblings are offspring of the same sire but different dams. Full-sib-half-sib design gives better estimation of genetic parameters in comparison to full-sib design, where each sire is mated to one unique dam (van der Waaij et al., 2019).

The paternal stock was gathered from the Bestico production line. Across three batches at random, 60 males were mated to two females each, resulting in 120 families in total. Virgin adults were paired up (one male:two females) in a petri-dish that had wet cotton to provide water, granulated sugar as feed and raw beef meat to lay eggs on (rearing room: 26 °C, relative humidity 50-60%). Females were isolated to their own individual petri-dish after 48 hours. The females were checked twice daily until six days after isolation, if no eggs were collected the mating was considered unsuccessful. When a decent sized clutch of eggs was found (>20 eggs), then the female was considered to have laid eggs and was removed. The clutch of eggs was divided into two groups, each placed in a new container providing 30g of larva feed (fish meal, soy meal and water, with ratio of 1:1:3) and reared in a warmer rearing room (28 °C, relative humidity 30-40%). It is difficult to identify the origin of the family if all half groups were reared

together, therefore the group factor (common environment) is nested within each family. Four days after the eggs were collected, larvae were moved to a petri-dish with 30g of new larva feed and three grams of sawdust. Maximum 50 to 60 individuals per family-group were kept and the remaining were discarded. Eight days after the eggs were collected, larvae were again moved to a new petri-dish but were only provided with eight grams of sawdust for pupation. On both 4th and 8th days, the family-group average weight was measured from ca. 20 random larvae per family-group. Individual weight could not be measured due to the difficulty to tag the larva. The petri-dishes were checked daily with intervals of 19 to 26 hours to record the pupated individual. Pupae were moved to the test tube with appropriate amounts of sawdust. Each test tube contained a maximum of 18 pupae from the same family-group that pupated on the same day. The test tubes were checked in the same daily time interval for adult emergence. The sex of the adult was recorded, then the flies were removed from the tube. The overview of the timeline can be seen in figure 1.

	Days																					
Action:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	...
Mate virgin parents (1 male:2 females)	x																						
Isolate mated females			x																				
Move eggs to new containers				x	x	x	x	x	x														
Move 4 days old larva to new feed								x	x	x	x	x	x										
Move 8 days old larva to saw dust												x	x	x	x	x	x						
Check pupation												x	x	x	x	x	x	x	x	x	x	x	x
Check emergence																x	x	x	x	x	x	x	x

Figure 1. Timeline of a batch.

2.2 Larval and pupal development time analysis

Genetic parameters of larval and pupal development time were estimated using the ASReml software (Gilmour et al., 2015). Two types of linear mixed model were run: sire model (SM) and animal model (AM). Each model has its advantages and disadvantages, both were computed to analyze which options are better for breeding of BDF. SM used the sire records to estimate the genetic parameters, whereas AM used all pedigree relations to estimate the genetic parameter. Generally, AM is superior to SM in accuracy of estimating results, but SM is sometimes used because less computation power is needed (Sun et al, 2009). In this report, the pedigree has only one generation, hence SM is comparable to AM. In addition, dam records were included in the SM as an extra random effect, which encapsulated and accounted for the maternal effect from the genetic parameters estimate. Maternal effect was corrected because it may have a noticeable impact on the development time of BDF, and result in

overestimation of the genetic parameter. Mousseau & Dingle (1991) reported that maternal effect has a significant effect on development time of multiple fly species.

The sire model was as follow:

$$y_{ijklm} = \mu + sex_i + batch_j + sire_k + dam_l + group_{m(l)} + \varepsilon_{ijklm}$$

y_{ijklm} was the trait value of the individual. μ was the population mean. sex_i represented linear fixed regression, with 0 as female, 1 as male, and individual that did not emerge (unknown sex) as 0.5, assuming 50:50 sex ratio. $batch_j$ was included as a fixed effect due to only having three batches. $sire_k$, dam_l, $group_{m(l)}$, and ε_{ijklm} were random effects. $sire_k$ was sire effect, and was used to estimate additive genetic variance. dam_l was dam effect and used to capture maternal effect. $group_{m(l)}$ was the group effect that was nested within the dam effect, and was used to estimate the common environmental variance. ε_{ijklm} was the residual effect. y was assumed to be normally distributed. $sire$, dam, $group$, and ε were assumed to be normally distributed and had equal variance.

Additive genetic variance (Va) was estimated by: $4 * Vsire$.
Heritability (h^2) was estimated by: $Va\ /(Vsire + Vdam + Vgroup + Vresidual)$
The contribution of the common environment to the phenotypic variation (c^2) was estimated by: $Vgroup\ /(Vsire + Vdam + Vgroup + Vresidual)$

Genetic, environmental, and phenotypic correlations were estimated with a bivariate model that includes both traits as y.
Additive genetic covariance was estimated by: $COVsire * 4$

The animal model was as follow:

$$y_{ijkl} = \mu + sex_i + batch_j + animal_k + group_{l(dam)} + \varepsilon_{ijkl}$$

y_{ijkl} was the trait value of the individual. μ was the population mean. sex_i represented linear fixed regression, with 0 as female, 1 as male, and individual that did not emerge (unknown sex) as 0.5, assuming 50:50 sex ratio. $batch_j$ was included as a fixed effect due to only having three batches. $animal_k$, $group_{l(dam)}$, and ε_{ijkl} were random effects. $animal_k$ was the animal effect and it used the pedigree data to estimate the additive genetic variance. $group_{l(dam)}$ was the group effect that was nested within the dam, and was used to estimate the common environmental variance. ε_{ijkl} was the residual effect. y was assumed to be normally distributed. $animal$, $group$, and ε were assumed to be normally distributed and had equal variance.

Heritability (h^2) was estimated by: $Vanimal\ /(Vanimal + Vgroup + Vresidual)$
The contribution of the common environment to the phenotypic variation (c^2) was estimated by: $Vgroup\ /(Vanimal + Vgroup + Vresidual)$

Genetic, environmental, and phenotypic correlations were estimated with a bivariate model that include both traits as y**.**

2.2 Additional data analysis: Larval weight and adult emergence rate

Tests performed in this section were run with R package (R Core Team, 2022).

Spearman rank correlation tests were used to find the correlation between: 4th-day/8th-day larval weight and larval/pupal development time, as well as between 4th-day/8th-day larval weight and emergence rate. Spearman rank correlation test was used because larval weight did not meet the normality assumptions. Larval weights, development time and emergence rate used the family-group average instead of individual larvae records.

Wilcoxon sum rank test (one-sided) was used to check the distribution of larval development time between emerged and un-emerged BDF. The Wilcoxon sum rank test was chosen because larval development time was not normally distributed under the emerged category. One-sided alternative hypothesis tested if the median of larval development time is higher in emerged BDF compared to unemerged BDF.

3. Results

3.1 Larval and pupal development time: Basic statistical analysis

There were in total 99 families/dams and 57 sires, of which provided 5509 larvae, and 3424 emerged adults. Larval and pupal development time were being measured on individual level, their descriptive statistics can be seen on table 1. Larval development time has higher standard deviation (SD) than pupal development time. It was shown that 95% of the individuals had a larval development time of eight to fifteen days, while 95% of individuals had a pupal development time of seven days.

Table 1. Descriptive statistics of larva and pupa development time.

Trait	Mean ± SD	Count	Min	Max
Larva development time (days)	11.09 ± 1.80	5509	8	18
Pupa development time (days)	7.04 ± 0.55	3424	5	10

Figure 2 shows the sex effect on the duration of larva/pupa stage. Males had slightly shorter larval development time than females, which was also supported by the LMM analysis that fixed male factor is -0.39 (p: <2.2e-16) days shorter. While it was shown that male had significantly longer pupal development time than females, however it was only 0.07 (p: 0.0001) days longer.

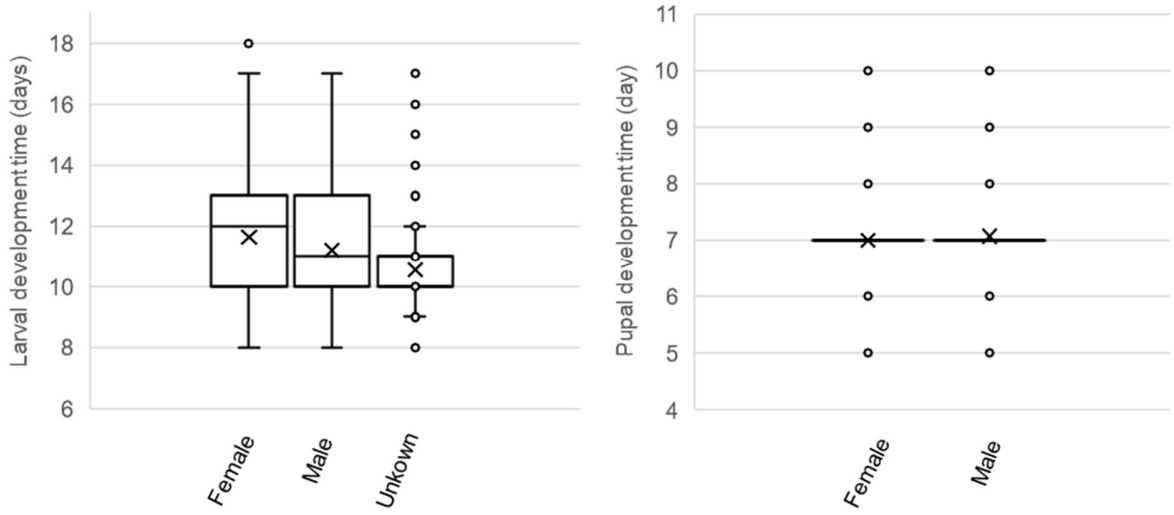

Figure 2. Trait distributions grouped by sex effect. The x represents the mean, the middle line of the box represents the median, boxes, and whiskers each represent a 25 percentile, o represents outliers.

Figure 3 shows the batch effect on the duration of larva/pupa stage. For both traits, batch one was significantly (p: <2.2e-16) differ from batch two and three. Batch one larval development time was four days shorter, while pupal development time was half a day shorter.

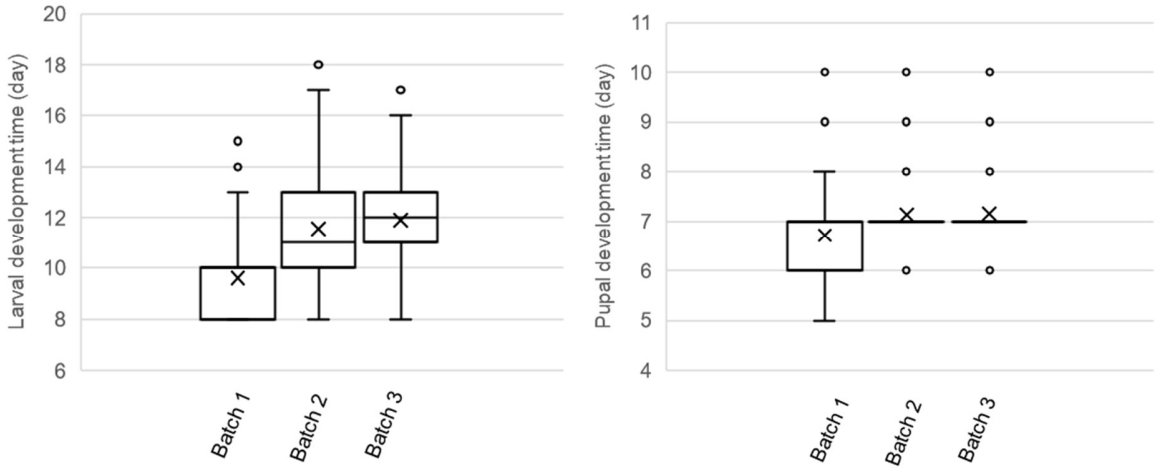

Figure 3. Trait distributions grouped by batch effect. The x represents the mean, the middle line of the box represents the median, boxes, and whiskers each represent a 25 percentile, o represents outliers.

3.2 Larval and pupal development time: Genetic parameters estimation

The estimated variances, heritability and common environmental effect of the larval development time and pupal development time can be found in table 2 and 3 respectively. Table 4 shows the genetic, environmental, and phenotypic correlations between the two development times.

The sire model shows that the heritability of larval development time is high (0.61, SE: 0.28), although the standard error is also high, it is still shown to be significantly different from zero. While the heritability of the pupal development time is much lower (0.05, SE: 0.08), and it cannot be said to significantly differ from zero. Common environmental effects of both traits are low (<0.1), with small standard errors. Additive genetic (0.12, SE: 0.62), environmental (0.14, SE:0.02) and phenotypic (0.17, SE: 0.03) correlations were all positive, but no strong correlation was found.

The Animal model shows a similar trend as the sire model, however the additive genetic variance of the larval development time is notably higher (2.10, SE: 0.33), which resulted in heritability closer to one (0.90, SE: 0.08). Additive genetic correlation (0.47, SE: 0.18) and environmental correlation (0.26, SE: 0.17) also had higher results compared to the sire model.

Table 2. Sire and animal model estimation of larval development time (days): additive genetic variance (Va), common environment variance (Vce), phenotypic variance (Vp), heritability (h^2) and common environmental effect (c^2) of the traits, with standard error (SE).

	Va ± SE	Vce ± SE	Vp ± SE	h^2± SE	c^2 ± SE
Sire model	1.36 ± 0.69	0.06 ± 0.02	2.26 ± 0.16	0.61 ± 0.28	0.02 ± 0.01
Animal model	2.10 ± 0.33	0.06 ± 0.02	2.34 ± 0.16	0.90 ± 0.08	0.02 ± 0.01

Table 3. Sire and animal model estimation of pupal development time (days): additive genetic variance (Va), common environment variance (Vce), phenotypic variance (Vp), heritability (h^2) and common environmental effect (c^2) of the traits, with standard error (SE).

	Va ± SE	Vce ± SE	Vp ± SE	h^2 ± SE	c^2 ± SE
Sire model	0.01 ± 0.02	0.03 ± 0.01	0.27 ± 0.01	0.05 ± 0.08	0.10 ± 0.02
Animal model	0.03 ± 0.01	0.03 ± 0.01	0.27 ± 0.01	0.09 ± 0.05	0.11 ± 0.02

Table 4. Sire and animal model output of: additive genetic correlation (rA), environmental correlation (rE) and phenotypic correlation (rP) between larval and pupal development time, with standard error (SE).

	rA ± SE	rE ± SE	rP ±SE
Sire model	0.12 ± 0.62	0.14 ± 0.02	0.17 ± 0.03

Animal model	0.47 ± 0.18	0.26 ± 0.17	0.17 ± 0.03

3.3 Additional data analysis: Larval weight

The weight of the larva was measured at four and eight days old. Due to difficulty to tag every larva, weight was measured on the family-group level. The descriptive statistics can be seen in table 5. In this report setup, the larva grew faster than in the commercial production setup. Some families reached the 20mg average larval weight threshold in four to six days, whereas in production it takes seven or more days. However, it did not change the minimum eight days requirement to start pupation.

Table 5. Descriptive statistics of family-group average larval weight.

Trait	Mean ± SD	Min	Max
4th day weight (mg)	16.07 ± 4.69	4	27
8th day weight (mg)	26.91 ± 2.26	18	33

Complete output of the Spearman rank correlation test can be seen in table 6. All Spearman rank correlation tests between larval weight traits and larval/pupal development time had a significant ($p<0.05$) correlation. Please note that these correlations were estimated on the average value per family-group. 4th-day weight had a weak negative correlation ($r<-0.17$) with both larval and pupal development time, and 8th-day weight had a moderate positive correlation ($0.25<r<0.5$) with larval and pupal development time. Negative correlation indicates that heavier larva at 4th day will have shorter development time, while positive correlation means that heavier 8th day old larva will take longer to develop. R-square value indicates how well the data fit the correlation trend. Low r-square ($R^2<0.23$) between all pairs showed that the overall trends were faint.

Table 6. Correlation (r) between larval weight (family-group average) and larval/pupal development time (family-group average). * indicates p value <0.05, ** indicates p value < 0.0001. Value in the [ ] indicates R-square.

	Larval development time (day)	Pupal development time (day)
4th day weight (mg)	-0.17* [0.03]	-0.05* [0.003]
8th day weight (mg)	0.48** [0.23]	0.30** [0.09]

3.4 Additional data analysis: Adult emergence

Wilcoxon sum rank test (one-sided) of larval development time between emerged and unemerged BDF showed significant results (p: <2.2e-16). This indicates that the median larval development time of emerging BDF is greater than median larval development

time of unemerged BDF. Which means that BDF with shorter larval development time tend to have lower emergence chance from pupa. A bar graph representation of the adult emergence rate by larval development time can be seen in figure 4.

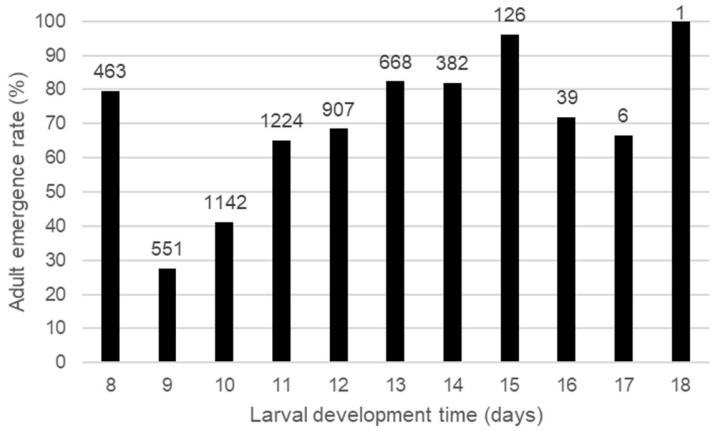

Figure 4. Adult emergence rate per larval development time.
The number on top of the bar represents the total sample size.

Spearman rank correlation tests between larval weight and emergence rate showed significant results ($p<0.0001$). Traits used in this analysis were at the family-group level. 4 days old larval weight had medium negative correlation with emergence rate, indicating that heavier larvae tend to have lower emergence. In contrast, 8 days old larval weight had medium positive correlation with emergence rate, which means that heavier larvae have higher emergence. However, both correlations had low r-square value ($R^2<0.15$), and this means that only a low percentage of records fit the trends. Output data can be seen on table 7.

Table 7. Correlation (r) between larval weight (family-group average) and emergence rate (family-group). * indicates p value <0.05, ** indicates p value < 0.0001. Value in the [ ] indicates R-square.

	Emergence rate
4th day weight (mg)	-0.38** [0.14]
8th day weight (mg)	0.25** [0.06]

4. Discussion

4.1 Model comparison

Animal model (AM) has been shown to be superior to the sire model (SM), as it provides more accuracy to the estimation. However, SM was also used widely because it required less computation power and was easier to implement (Sun et al, 2009). Both AM and SM have their own usage depending on the breeding design and available time/budget. SM is more preferred under the setup of this report, because the pedigree contains only one generation, thus easier to implement and run. However, Sun et al (2009) suggested that the accuracy of SM decreases if there is non-random mating in the record, which would be the case for a breeding program. AM is a better option if selection has been done to the BDF, and/or the data consists of records from multiple generations, since AM takes account of all the pedigree relationships across generations.

In this report, the AM estimates had higher additive genetic variance, heritability, and genetic correlation, in comparison to SM. The SM applied here used sire effect to estimate the genetic variance, while including extra dam effect that accounted for the maternal effect (common environment effect due to the dam). The AM applied here used animal effect (pedigree with both sire and dam) to estimate the genetic variance, however, it did not include a separate dam effect that accounted for the maternal effect. This shows that: 1) The maternal effect has a noticeable impact on the phenotypic variance of the BDF's development time traits, and should be accounted for in the model. This is in line with Mousseau & Dingle (1991) report, which stated that maternal effects have a significant impact on the development time of several fly species. 2) Due to the absence of an extra dam effect in AM, the higher estimation is likely due to overestimation of the additive genetic variance. By including a dam random effect (not linked to pedigree) in the AM, the results would be more accurate. This is because the animal effect captures the genetic variance from the dam side, and the dam effect captures the maternal effect. For the data used in this report, the addition of dam effect would reduce the AM to SM, and result in a similar estimation. This is because the pedigree has only one generation, and no half-sib records for the dam. Therefore the animal effect would estimate the genetic parameter using primarily sire information, while dam effect captured variation due to the dam, including maternal effect.

AM ran in this report was not corrected for the maternal common environmental effect, while SM was corrected, therefore the results of SM will be used for the remainder of this report.

4.2 Larval development time

The larval development time had the mean of 11 days with a standard deviation (SD) of 1.8, which is similar to the result of Lefebvre & Pasquerault (2004) (mean: 6.6-9.5 days, SD: 0.69-1.45) and Hogsette & Washington (1995) (mean: 8.5-11 days, SD: 1.3-3). Both of these papers studied the effect of the environment on the larval development time. Lefebvre & Pasquerault (2004) showed that the mean and SD are both inversely proportional to the rearing temperature. As the temperature increased from 24C toward 30C, the mean shortened by 3 days and SD decreased by 0.76. Hogsette & Washington (1995) showed that by reducing the protein in the larva meal from 50% to 14% the mean development time decreased by 1.5 days and SD decreased by 1. This implies that by optimizing the rearing conditions, a sizable part of the variations of larval development time can be reduced.

Common environmental effects of BDF's larval development time is low (c^2: 0.02; standard error, SE: 0.01), especially in comparison to that of egg to prepupal development time of black soldier fly (c^2: 0.25; SE: 0.06; Bouwman et al., 2022). However, Bouwman et al. (2022) did not include half sib in their design, therefore the c^2 could also consist of the maternal effect, whereas the maternal effect of BDF is captured by the dam effect. Low c^2 indicates that the difference of phenotype between two groups of the same family is small, and it can be due to two reasons, 1) there was little difference under the setup of this experiment between the two environments each family was house in (e.g. temperature, humidity, feed composition, larva density), and/or 2) the common environmental effect on the phenotypic variance is low.

The narrow-sense heritability of the larval development time is 0.61 (SE: 0.28), which is relatively high in comparison to insects of the same order. Lehmann et al. (2006) showed that African malaria mosquito (*Anopheles gambiae*) had broad-sense heritability of 0.12-0.19 for larval development time, and Miyatake (1995) showed that melon fly (*Bactrocera cucurbitae*) had realized-heritability of 0.03-0.37 for its larval development time. This could indicate that the BDF's heritability is overestimated, as suggested by the high SE. To test it, further independent study should be performed. However, even when taking the lower bound of the 95% confidence interval, the heritability (h^2: 0.34) is still considerably high for selection. This shows that BDF's larval development time is potentially a good trait to be selected for, as higher heritability means the response to selection could also be higher.

4.3 Pupal development time

In comparison, the pupal development time had a smaller mean and standard deviation (mean: 7, SD: 0.55) than larval development time. This trend was also shown in common fruit fly (*Drosophila melanogaster*; Bakker & Nelissen, 1963) and European

drone fly (*Eristalls arbustorum*; Ottenheim et al., 1996). However, D'Almeida et al. (1999) had opposite trends in their experiment with BDF, which showed that the pupal development time lasted longer with higher SD than larval development time. A potential explanation to D'Almeida et al. (1999) results could be that they used a high protein diet. Which indicates that to some extent the environmental effect during the larva stage could have had a substantial effect on the pupal development time.

The heritability of pupal development time is 0.05 (SE: 0.08), and it cannot be said to be significantly differ from zero. The low variation and the low heritability show that pupal development time has less potential for selection than the larval development time. This finding is in agreement with Miyatake (1995) report, which found that the melon flies that were selected for longer overall development duration (egg to adult emergence), their larval development time had a more drastic response to selection compared to pupal development time.

4.4 Correlation

The phenotypic correlation (0.17, SE: 0.03), environmental correlation (0.14, SE: 0.02) and genetic correlation (0.12, SE: 0.62) between larval and pupal development time were all positive. This is desirable, because the aim of the selection is to have both traits improve in the same direction toward shorter development time, but the correlations were considerably low. As a reference, the common fruit fly has the phenotypic correlation of 0.4-0.5 for larva and pupal development time, and has shown to have noticeable effect on the response of one trait when selection was done to the other trait (Bakker & Nelissen, 1963). The low heritability of pupal development time and the low positive genetic correlation means that pupal development time is not a priority trait to be selected for. However, there are other traits that may be important to investigate for its correlation with the larval development time. Three of these traits are: 1) adult emergence rate, 2) sex ratio, and 3) pre-/oviposition duration.

In this report, a trend was shown between shorter larval development time and lower adult emergence rate. Prasad et al. (2001) found a similar trend that common fruit fly has positive correlation between faster development time and higher premature mortality rate. BDF is sold as a live product, therefore the emergence rate is tied to the product quality. Having a higher adult emergence rate will be more desirable.

Sex effect is significant on the larval development time, male had shorter mean larval development time than females, which is aligned with the result from Johnson & Venard (1957). Selection based on shorter development time could result in more males than females in the emerging adult, and thus may lower the number of available offspring for

the next generation. As a biocontrol product, it is also more desirable to have a higher female to male ratio, as they can lay more larvae which are the main controlling agents.

Pre-oviposition duration is the time between the emergence from pupa (adult stage) to laying first clutch of egg, and oviposition duration is the time between the first and last egg clutch. Both traits are female specific and can influence the synchronization of the BDF's life cycle. Observationally, the pre-oviposition lasted two to eight days after mating, and peaked at four days. Oviposition duration was not known, but it was seen that each female had different egg laying behavior. Some laid eggs in a big clutch and rested a few days until the next clutch, while some laid a continuous small number of eggs throughout the day. Mcinnis et al. (1983) found in screwworm (*Cochliomyia hominivorax*) that smaller individuals correlate with shorter development time, which also results in shorter pre-oviposition and oviposition duration. However, they also showed that the duration of oviposition did not correlate with the number of eggs a female can lay in its lifetime. The synchronization of the BDF production can be further improved through optimizing the pre-/oviposition duration traits, but a more in-depth study is needed.

Overall, it seems that faster larval development time may be tied to undesirable results in some other important traits. It is recommended to have stabilizing selection toward a mid-fast larval development time, instead of directional selection toward faster larva. As it would still improve the syncing of the production cycle without introducing new potential issues.

The correlations between larval weight and larval-pupal development time/adult emergence rate had low r-square (<0.23). R-square represents the proportion of data that fit the trend, which means that larval weight is not a good predictor for larval-pupal development time and adult emergence rate in this setup. Larval weight is not a key trait that affects the biological control effectiveness of BDF, however it is an easy to measure trait that can be used to monitor the general vitality of BDF in production.

4.5 Breeding program recommendation

There are three options for the breeding program of BDF: 1) mass selection, 2) Individual selection (with pedigree information), and 3) between-family selection.

In mass selection, the individuals with the best phenotype are selected as parents. Mass selection requires the least labor, because the pedigree record is not needed for estimating the breeding value (BV, genetic potential), thus BDF of the same stage can be reared together in one place without tracking. This allows more BDF to be reared at once, which means that more individuals can be selected as parents, while keeping the

same selection intensity. Higher number of selected parents would slow down the increase of inbreeding, because the number of selected parents and increased rate of inbreeding are inversely correlated (Oldenbroek & van der Waaij, 2015). However, the lack of pedigree records means that the accuracy of estimated BV is also the lowest among the methods described here, as mass selection only uses the individual's own performance to estimate the BV, and no other information. Lower accuracy may lead to less response to selection, but accuracy also has a direct relationship with heritability, where higher heritability will lead to higher accuracy (Oldenbroek & van der Waaij, 2015). Mass selection is a viable option for BDF's larval development time, because it has medium to high heritability, therefore the accuracy would be sufficient. It is important to note that as selection continues, the heritability may decrease, which would also decrease the effectiveness of mass selection.

Individual selection (with pedigree information) and between-family selection both used the individual's own performance and pedigree information to estimate BV. The accuracy of the BV is higher than mass selection, because it includes the family relationship in the model for BV estimation, thus it may lead to higher response to selection (Oldenbroek & van der Waaij, 2015). However, BDF needs to be individually tracked for the pedigree, which means the labor requirement is higher and less candidates can be kept. With a smaller number of selected parents, the increase of inbreeding will be higher.

Individual selection (with pedigree information) uses the BDFs with the optimal BV from the total candidates as parents, which means that the response to selection would be the highest. However, this selection method also has the highest increase of inbreeding per generation, because individuals from the same family tend to have a similar BV, thus relatives are more likely to mate. This is not sustainable in the long term, due to the negative effect of the inbreeding depression. It was found in black soldier fly that inbreeding depression lowered the adult emergence rate (Rhode et al., 2020) adult longevity and fecundity (Badenhorst, 2017).

Between-family selection takes the best BDFs from each family, and mates the individuals from different families together. The response to selection may be less than individual selection (with pedigree information), because families with lower BV would be included. However, between-family selection is better than individual selection in the long term, because the increase of inbreeding is also lower, when the family mating is controlled and done in a circular manner (Oldenbroek & van der Waaij, 2015). Between-family selection would have more accurate BV than mass selection, but it also has a higher rate of inbreeding, and more labor is needed. Which of the two is a better option depends on the heritability of the traits, the budget/time and the goal of the selection.

In conclusion, all three breeding program options are viable for selection of BDF for optimal larval development time. Mass selection is recommended if the trait's heritability is high, labor is limited and/or if inbreeding is a major concern. Individual selection (with pedigree information) would have the highest response to selection, but the inbreeding rate would increase the fastest, which is not ideal for a long-term breeding program. Between-family selection is a good middle ground that may result in higher selection response than mass selection, while keeping the increase of inbreeding lower than individual selection. However, between-family selection will require more labor than mass selection.

5. Conclusions

Synchronizing and shortening the BDF's larval development time is possible through the means of breeding, while BDF's pupal development time does not have enough genetic potential to be selected on. The high heritability of larval development time means that mass selection or selection with pedigree are both feasible breeding design choices. It is recommended to select larva with a mid-fast development time, instead of shortening toward fastest development time, due to the possibility of short larval development negatively impacting other vital traits. Further research could be done to investigate the correlation between larval development time and 1) adult emergence rate, 2) sex ratio, 3) pre-/oviposition duration. These traits are equally important to the effectiveness of biological control, therefore it is recommendable that they are taken into consideration when performing selection.

Acknowledgment

I would like to thank my supervisors, Aniek Bouwman and Esther Ellen, for giving me guidance on writing this report, especially on developing the model and advice on ASReml. I would not have finished my thesis in time without the help.

I am grateful for the opportunity Bestico gave. I could not perform this research without the help of Jeroen van Schelt, who coordinated my work in Bestico and gave helpful feedback during data collecting. Lastly, I would like to thank the R&D crews from Bestico who helped me troubleshoot practical issues and gave me inspiration.

References

Badenhorst, R. (2017). *Genetic Diversity in a Commercial Black Soldier Fly, Hermetia illucens (Diptera: Stratiomydiae), population*. [Doctoral dissertation, thesis master of agriscience science.–Stellenbosch University] http://scholar.sun.ac.za/handle/10019.1/101182

Bakker, K., & Nelissen, F. X. (1963). ON THE RELATIONS BETWEEN THE DURATION OF THE LARVAL AND PUPAL PERIOD, WEIGHT AND DIURNAL RHYTHM IN EMERGENCE IN DROSOPHILA MELANOGASTER. *Entomologia Experimentalis et Applicata*, *6*(1), 37–52. https://doi.org/10.1111/J.1570-7458.1963.TB00601.X

Bouwman, A. C., Nugroho, J. E., Wongso, D., van Schelt, J., Pannebakker, B. A., Zwaan, B. J., & Ellen, E. D. (2022). Genetic parameters of Black Soldier Flies estimated in full sib design. Proceeding of the 12th World Congress on Genetics Applied to Livestock Production, Rotterdam, The Netherlands

Cortinhas, L. B., Mendonça, P. M., Barbosa, R. R., & Queiroz, M. M. de C. (2016). Ultrastructure of immature stages of the black dump fly: Ophyra aenescens (Wiedemann, 1830) (Diptera: Muscidae: Azeliinae). *Acta Tropica*, *158*, 125–129. https://doi.org/10.1016/J.ACTATROPICA.2016.02.022

D'Almeida, J. M., Borges, C., & Gonçalves, C. A. (1999). Desenvolvimento Pós-embrionário de Ophyra aenescens (Wiedemann, 1830) (Diptera: Muscidae) em Diferentes Dietas, sob Condições de Laboratório. *Memórias Do Instituto Oswaldo Cruz*, *94*(1), 123–126. https://doi.org/10.1590/S0074-02761999000100024

Facchini, E., Shrestha, K., van den Boer, E., Junes, P., Sader, G., Peeters, K., & Schmitt, E. (2022). Long-Term Artificial Selection for Increased Larval Body Weight of Hermetia illucens in Industrial Settings. *Frontiers in Genetics*, *13*. https://doi.org/10.3389/fgene.2022.865490

Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J. and Thompson, R. (2015). ASReml User Guide Release 4.2 Structural Specification, VSN International Ltd, Hemel Hempstead, HP1 1ES, UK www.vsni.co.uk

Hogsette, J. A., & Washington, F. (1995). Quantitative Mass Production of Hydrotaea aenescens (Diptera: Muscidae). In *J. Econ. Entomol* (Vol. 88, Issue 5). https://academic.oup.com/jee/article/88/5/1238/2216409

Hogsette, J. A., Farkas, R., & Coler, R. R. (2002). Development of Hydrotaea aenescens (Diptera: Muscidae) in Manure of Unweaned Dairy Calves and Lactating Cows. *Journal of Economic Entomology*, *95*(2), 527–530. https://doi.org/10.1603/0022-0493-95.2.527

Hogsette, J. A., & Jacobs, R. D. (2003). The Black Dump Fly: A larval predator of house flies. *Publications from USDA-ARS / UNL Faculty. 994.* . https://digitalcommons.unl.edu/usdaarsfacpub/994

Johnson, W. T., & Venard, C. E. (1957). Observations on the Biology and Morphology of Ophyra Aenescens (Diptera: Muscidae). *The Ohio Journal of Science*, *57*(1), 21–26. https://kb.osu.edu/handle/1811/4413

Lefebvre, F., & Pasquerault, T. (2004). Temperature-dependent development of Ophyra aenescens (Wiedemann, 1830) and Ophyra capensis (Wiedemann, 1818) (Diptera, Muscidae). *Forensic Science International*, *139*(1), 75–79. https://doi.org/10.1016/J.FORSCIINT.2003.10.014

Lehmann, T., Dalton, R., Kim, E. H., Dahl, E., Diabate, A., Dabire, R., & Dujardin, J. P. (2006). Genetic contribution to variation in larval development time, adult size, and longevity of starved adults of Anopheles gambiae. *Infection, Genetics and Evolution*, *6*(5), 410–416. https://doi.org/10.1016/j.meegid.2006.01.007

Mcinnis, D. O., Wendel, L. E., & Whitten, C. J. (1983). Directional Selection and Heritability for Pupal Weight in the Screwworm, Cochliomyia hominivorax (Diptera: Calliphoridae). *Annals of the Entomological Society of America*, *76*(1), 30–36. https://doi.org/10.1093/AESA/76.1.30

Miyatake, T. (1995). Two-Way Artificial Selection for Developmental Period in Bactrocera cucurbitae (Diptera: Tephritidae). *Japan Ann. Entomol. Soc. Am*, *88*(6), 848–855. https://academic.oup.com/aesa/article/88/6/848/163023

Mousseau, T. A., & Dingle, H. (1991). Maternal effects in insects: examples, constraints, and geographic variation. *The unity of evolutionary biology, 2,* 745-761.

Nolan, M. P., & Kissam, J. B. (1987). Nuisance Potential of a Dump Fly, Ophyra aenescens (Diptera: Muscidae), Breeding at Poultry Farms. *Environmental Entomology*, *16*(3), 828–831. https://doi.org/10.1093/EE/16.3.828

Oldenbroek, K., & van der Waaij, L. (2015). *Textbook Animal Breeding and Genetics for BSc students*. Centre for Genetic Resources The Netherlands and Animal Breeding and Genomics Centre. https://wiki.groenkennisnet.nl/display/TAB/

Ottenheim, M. M., Volmer, A. D., & Holloway, G. J. (1996). The genetics of phenotypic plasticity in adult abdominal colour pattern of Eristalis arbustorum (Diptera: Syrphidae). *Heredity 1996 77:5*, *77*(5), 493–499. https://doi.org/10.1038/hdy.1996.176

Prasad, N. G., Shakarad, M., Anitha, D., Rajamani, M., & Joshi, A. (2001). Correlated responses to selection for faster development and early reproduction in Drosophila: the evolution of larval traits. *Evolution; International Journal of Organic Evolution*, *55*(7), 1363–1372. https://doi.org/10.1111/J.0014-3820.2001.TB00658.X

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. "R version 4.1.3 (2022-03-10)" https://www.R-project.org/

Rhode, C., Badenhorst, R., Hull, K. L., Greenwood, M. P., Bester-van der Merwe, A. E., Andere, A. A., & Richards, C. (2020). Genetic and phenotypic consequences of early domestication in black soldier flies (Hermetia illucens). *Animal Genetics, 51*(5), 752-762. https://doi:10.1111/age.12961

Sun, C., Madsen, P., Nielsen, U. S., Zhang, Y., Lund, M. S., & Su, G. (2009). Comparison between a sire model and an animal model for genetic evaluation of fertility traits in Danish Holstein population. *Journal of Dairy Science*, *92*(8), 4063–4071. https://doi.org/10.3168/jds.2008-1918

Appendix I: Data Management Plan

Data management plan belonging to the MSc thesis performed at the Animal Breeding and Genomics Group by David Wongso, completed in January 2023.

Agreements

1. The data used in this thesis project have been described in this document and have been stored in a systematic manner (at least in separate folders for all sections as described below). Data includes all data as mentioned in the results section of your report.
2. The data management plan has been discussed with the MSc thesis supervisor and he/she has agreed on the location for data storage.
3. In case of confidentiality, contact details of the responsible person from the company/institution that has ownership of the data are mentioned in this document.
4. **The data can be found in/on/through Aniek Bouwman (aniek.bouwman@wur.nl). Location: W:\ASG\WLR_Genomica\Projects\Sustainable_Breeding\Bestico**

Section A: Raw data

Created in (year, month)	Created in (year, month)	Remark
Raw collected data.xlsx	2022, Sept	Original raw data
bdf_data_final.xlsx	2022, Nov	Organized data for R-script
bdf_data_fam.xlsx	2022, Dec	Organized data for R-script
bdffinal.csv	2022, Nov	Organized data for ASReml
bdfped.ped	2022, Nov	Organized data for ASReml
bdfped2.ped	2022, Nov	Organized data for ASReml

Comments: All data collected by David Wongso in Bestico BV (Berkel en Rodenrijs, the Netherlands) from 2022 Sept to 2022 Nov, and processed for analysis from 2022 Nov to 2022 Dec.

Section B: Data analysis (e.g. script files)

File names	Created in (year, month)	Remark
bdf_statistic.R	2022, Nov	R script to analyze basic statistic, assumption and pre-run of sire model: larval/pupal development time
Larva (SM)\model.as	2022, Nov	ASRmel script to analyze sire model: larval development time

Pupa (SM)\model.as	2022, Nov	ASRmel script to analyze sire model: pupal development time
Bivariate (SM)\model.as	2022, Nov	ASRmel script to analyze animal model: larval-pupal development time correlation
Larva (AM)\model.as	2022, Dec	ASRmel script to analyze animal model: larval development time
Pupa (AM)\model.as	2022, Dec	ASRmel script to analyze animal model: pupal development time
Bivariate (AM)\model.as	2022, Dec	ASRmel script to analyze animal model: larval-pupal development time correlation

Section C: Final data

File names	Created in (year, month)	Remark
BDF_figures.xlsx	2022, Nov	All figures in the report
BDF_tables.xlsx	2022, Nov	All tables in the report
Larva (SM)\model.asr	2022, Nov	Data used in table 2
Larva (SM)\model.pvc	2022, Nov	Data used in table 2
Pupa (SM)\model.asr	2022, Nov	Data used in table 3
Pupa (SM)\model.pvc	2022, Nov	Data used in table 3
Bivariate (SM)\model.asr	2022, Nov	Data used in table 4
Bivariate (SM)\model.pvc	2022, Nov	Data used in table 4
Larva (AM)\model.asr	2022, Dec	Data used in table 2
Larva (AM)\model.pvc	2022, Dec	Data used in table 2
Pupa (AM)\model.asr	2022, Dec	Data used in table 3
Pupa (AM)\model.pvc	2022, Dec	Data used in table 3
Bivariate (AM)\model.asr	2022, Dec	Data used in table 4
Bivariate (AM)\model.pvc	2022, Dec	Data used in table 4