1 AIM

To determine whether the alteration of *Pinus radiata* [redacted] genes has affected the height and trunk diameter of the trees or the survival and growth of the geometrid *Pseudocoremia suavis* feeding on the modified material.

2 MATERIALS AND METHODS

2.1 Trees

Trees were of five clonal lines (Table 1): one unmodified control (Line 126) that had been through embryogenesis to ensure that its treatment was as similar as possible to that of the transgenic lines; two transgenic control lines from distinct transformation events that contained the empty vector ([redacted] Lines 103 and 104); and two distinct lines containing the vector expressing modified [redacted] (Lines 20 and 135). Plants were all of the same basic genotype.

Table 1: The five lines of *Pinus radiata* trees used in this work.

Line 126	**isogenic control**
Line 103	**contains the vector but no** [redacted]
Line 104	**contains the vector but no** [redacted]
Line 20	[redacted]
Line 135	[redacted]

The trees were maintained and grown in a PC2 level containment glasshouse at the [redacted] Scion provided nine trees of Line 135 and 10 trees of all other lines for this project. To ensure the degree of damage due to needle harvesting was equivalent across all lines, nine trees of each line were used in insect assays. This would ensure that the up-regulation of inducible compounds with potential anti-insect feeding properties, such as protease inhibitors, would be equivalent in all lines.

At the beginning of the two bioassay replicates in May 2014 and in May 2015, the height of each tree was measured as the distance from the soil to the tip of highest needles on the central leader, and the trunk diameter of each tree was measured using digital callipers at 5 cm above the soil. Two diameter measurements were taken for every trunk and the average value used.

2.2 Insect assay

Neonate larvae of the common forest looper, *Pseudocoremia suavis*, were obtained from a laboratory colony [redacted] Larvae were assigned randomly to be fed exclusively on needles from one of the five pine lines and reared individually in pots from neonate larva to pupation and then to adult (moth) emergence. Pine needles were harvested evenly from all plants in each pine line and sterilised in 0.5% hypochlorite solution, rinsed and dried before each feeding. Larvae were monitored daily for death, pupation, and emergence as adults. The insects were weighed *en masse* as neonates at the beginning of the assays and then on nine additional weigh days throughout the larval stage. Each individual was also weighed on the day it became a pupa.

A total of 45 larvae were reared on each pine line, i.e. 20 in the first replicate in June 2014 and 25 in the second replicate in June 2015. Numbers in the two replicate bioassays were limited by the size of the trees and thus the availability of pine needle material for feeding. Likewise the scheduling of the assays depended upon sufficient growth of the trees to provide the needle material required.

2.3 Data analysis

Analysis of variance (ANOVA) was used to detect any differences among treatments in the height and girth of trees, larval and pupal weight and absolute growth rate using Minitab® (Minitab Statistical Software (2010), version 16.1.1. State College, PA: Minitab, Inc. www.minitab.com). Both Minitab® and G*Power (version 3.1.9.2, http://www.gpower.hhu.de/en.html, Heinrich Heine Universität Düsseldorf) were used for power analysis. To reduce the likelihood of spurious differences being detected when multiple comparisons were made, the Bonferroni multiple comparison procedure was used. Treatment effects on survival, time to pupation, duration of pupal stage and total time from neonate larva to moth emergence, were tested using the Log-Rank and Wilcoxon tests to compare Kaplan-Meier survival curves. Data from the two replicates were pooled for the Kaplan-Meier curves. For relative growth rates, the Kruskal-Wallis non-parametric test was used, as residuals were not normally distributed even after transformation.

3 RESULTS

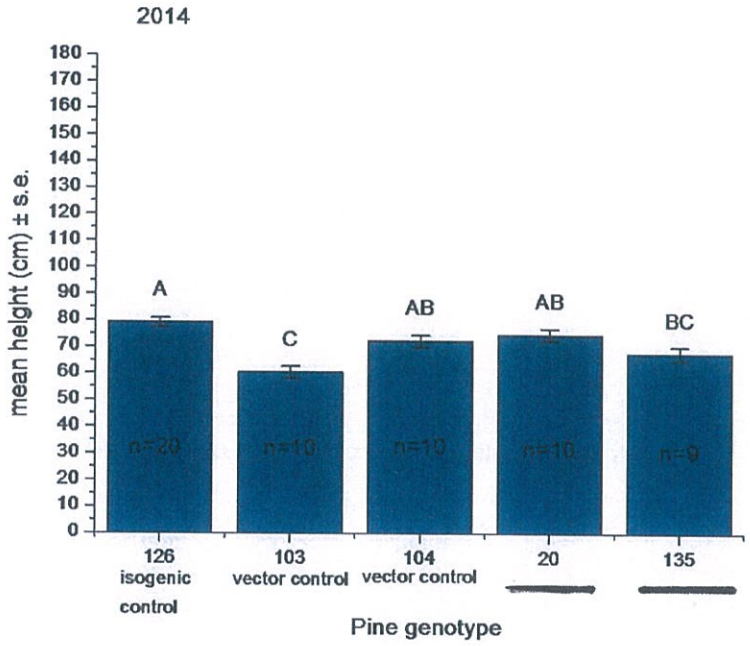
3.1 Tree phenotypes

Significant differences in mean height and mean trunk diameter were found among the different pine lines (Figures 1 and 2). At the first measurement the isogenic control line trees were on average the tallest, being significantly different from the shortest lines which were the vector-only control Line 103 and the [redacted] Line 135. At the second measurement the [redacted] Line 20 trees were on average tallest but not significantly different from those of the isogenic control. The shortest line, Line 103, was still significantly different from the isogenic control.

Trees of the isogenic control line had the greatest mean diameter at both dates. The mean diameter was significantly different from those in all other pine lines on the first date, but a year later the control mean was significantly different only to the lowest mean which was now Line 103 (previously Line 135).

The small number of trees of each line (20 for isogenic controls, nine for Line 135 and 10 for the other lines) means that the ANOVAs are likely to have low power and a relatively high probability of Type II error, i.e. incorrect acceptance of the null hypothesis that means do not differ. Thus additional differences could also exist. It would be of interest to determine whether phenotypic differences among pine lines persist as the trees grow.

A.



B.

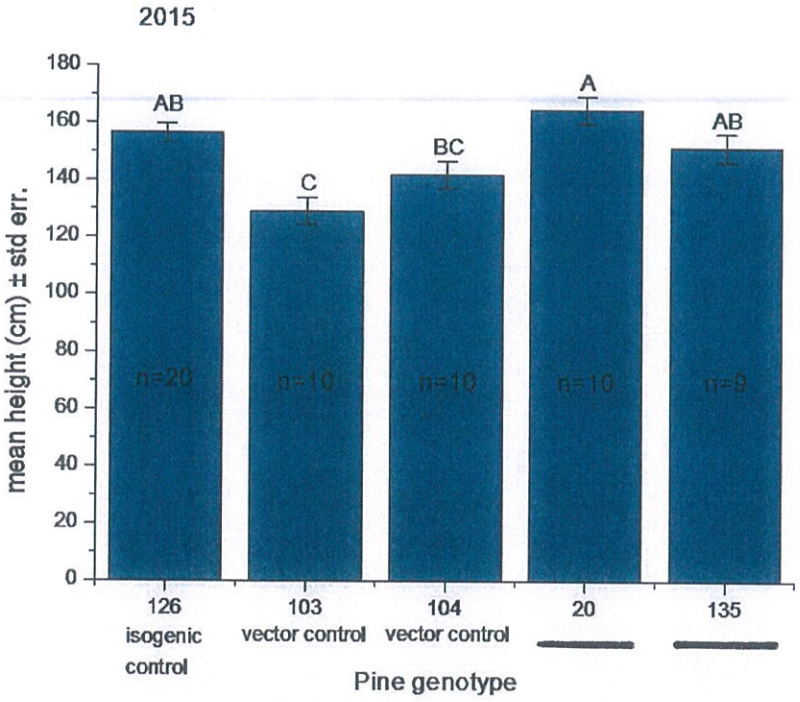
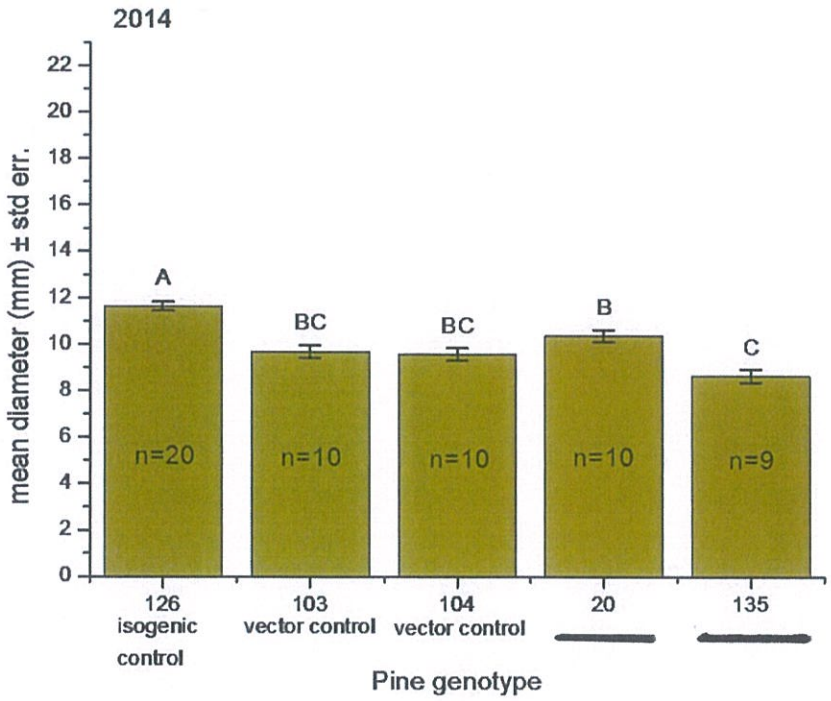


Figure 1. Mean *Pinus radiata* tree heights (a) May 2014 and (b) May 2015. Means sharing a letter are not significantly different (Bonferroni multiple comparison procedure at the 0.05 family level of significance).

A.



B.

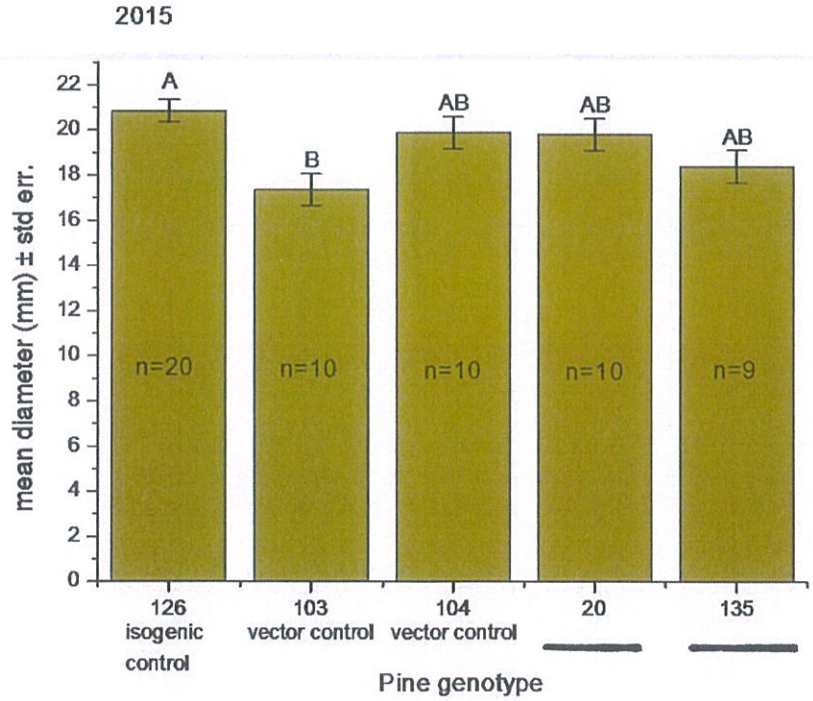


Figure 2. Mean diameter of *Pinus radiata* tree trunks measured 5 cm from the soil (a) May 2014 and (b) May 2015. Means sharing a letter are not significantly different (Bonferroni multiple comparison procedure at the 0.05 family level of significance).

3.2 Survival and development of *P. suavis*

3.2.1 Larval survival

Few larvae died in any of the treatments and no significant differences were found among the Kaplan-Meier survival curves of larvae fed on the five pine lines (Log-Rank $\chi^2 = 7.11115$, d.f. = 4, P = 0.130; Wilcoxon $\chi^2 = 7.42006$, d.f. = 4, P = 0.115, n = 45 per treatment) (Figure 3). *Post hoc* power analysis indicated that with this sample size the statistical power of the test to find a difference where a population difference of 10% actually exists was approximately 50%. This means that the probability of Type II error (incorrectly accepting the null hypothesis) if a difference of 10% exists is about 50%. A total sample size of 90-300 per treatment may be required to give 80% statistical power.

3.2.2 Larval growth

No significant differences in mean larval weights were found among pine lines on each weigh day, except for a difference between the two vector control treatments (Lines 103 and 104) on Day 7 only ($F_{4,177} = 2.70$, P = 0.032) (Figure 4). Differences were not tested after Day 35 because of the small numbers of larvae remaining, as most had become pre-pupae or pupated by this time.

The biggest variation in relative growth rate (RGR) among individuals was in the first seven days and so RGR for this time period was calculated and analysed. A Kruskal-Wallis analysis found no differences among the pine lines (H = 7.74, d.f. = 4, P = 0.102). RGR in this period was higher in replicate 1 (H = 12.45, d.f. = 1, P < 0.001) but there was no gender difference (H = 3.84, d.f. = 2, P = 0.147).

3.2.3 Development times

Larval stage duration

Kaplan-Meier curves of 'time from egg hatch to pupation' did not significantly differ among *P. suavis* fed on the different pine lines (Log-Rank test: $\chi^2 = 3.86087$, d.f. = 4, P = 0.425; Wilcoxon: $\chi^2 = 4.12848$, d.f. = 4, P = 0.389). Sample sizes were: line 126 n = 40; line 103 n = 37; line 104 n = 33; line 20 n = 40; line 135 n = 39 (Figure 5).

Pupal stage duration

Kaplan-Meier curves of 'days as pupa' did not differ among *P. suavis* fed on the different pine lines (Log-Rank test: $\chi^2 = 5.41758$, d.f. = 4, P = 0.247; Wilcoxon: $\chi^2 = 3.63662$, d.f. = 4, P = 0.457). Sample sizes were: line 126 n = 34; line 103 n = 29; line 104 n = 28; line 20 n = 27; line 135 n = 34 (Figure 6).

Total time to adult emergence

No differences between Kaplan-Meier curves of 'days to emergence as adults' were found among pine lines (Log-Rank test: $\chi^2 = 6.50796$, d.f. = 4, P = 0.164; Wilcoxon: $\chi^2 = 7.82581$, d.f. = 4, P = 0.098). Sample sizes were: line 126 n = 34; line 103 n = 29; line 104 n = 28; line 20 n = 27; line 135 n = 34 (Figure 7).

The absence of any differences in development times suggests that none of the pine lines provided a significantly poorer or better quality food source for the larvae. However, for these parameters, the sample sizes did not give high power in the statistical tests, and undetected differences may yet exist. Only non-parametric survival analyses were suitable for these duration data. The power of the tests to detect a difference if a population difference of 10% exists is likely to be less than 40% for the larval, pupal, and total development duration.

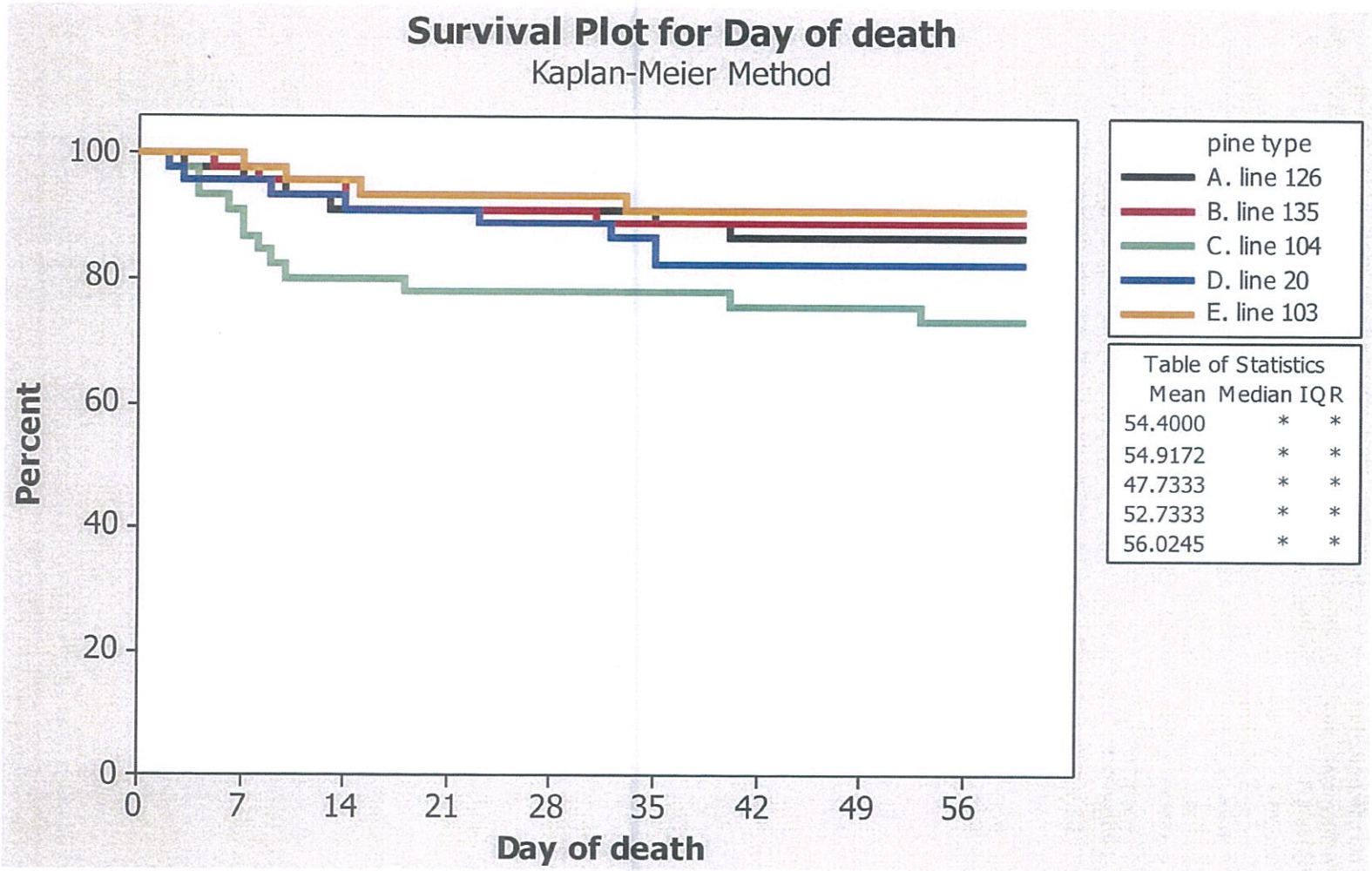


Figure 3. ***Pseudocoremia suavis*** **survival from egg hatch to pupation on different** ***Pinus radiata*** **lines. Data from the two replicates were pooled.**

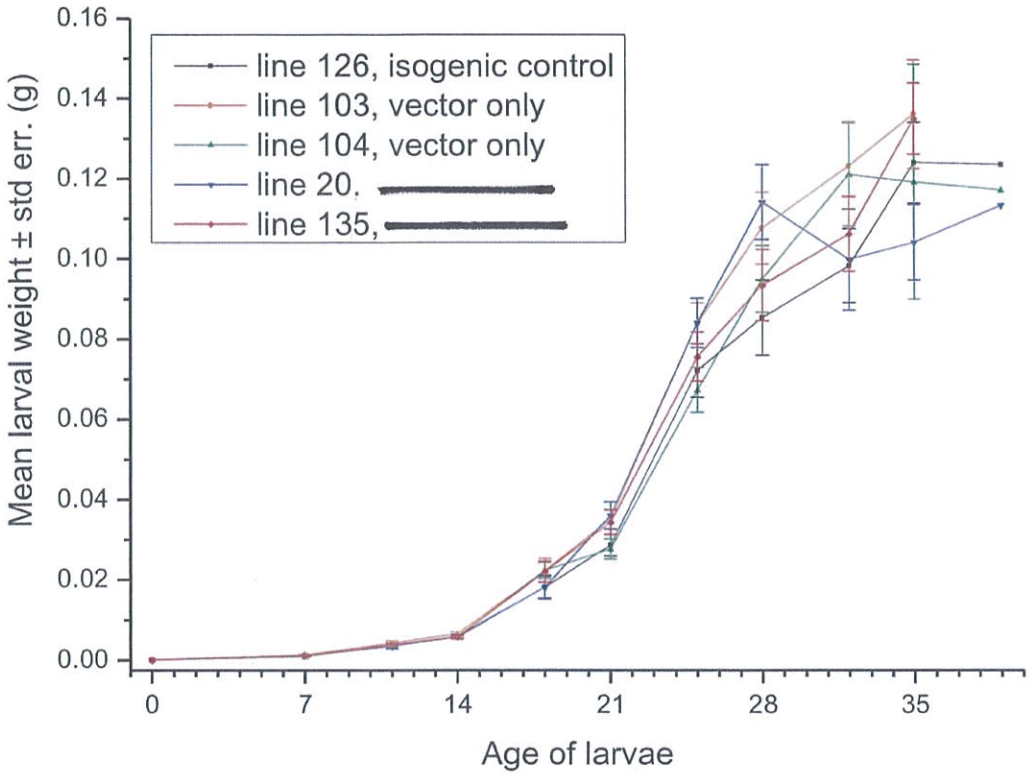


Figure 4. Larval growth, i.e. mean larval weights through time of *Pseudocoremia suavis* fed on five different *Pinus radiata* lines. Raw means and standard errors of the mean are shown.

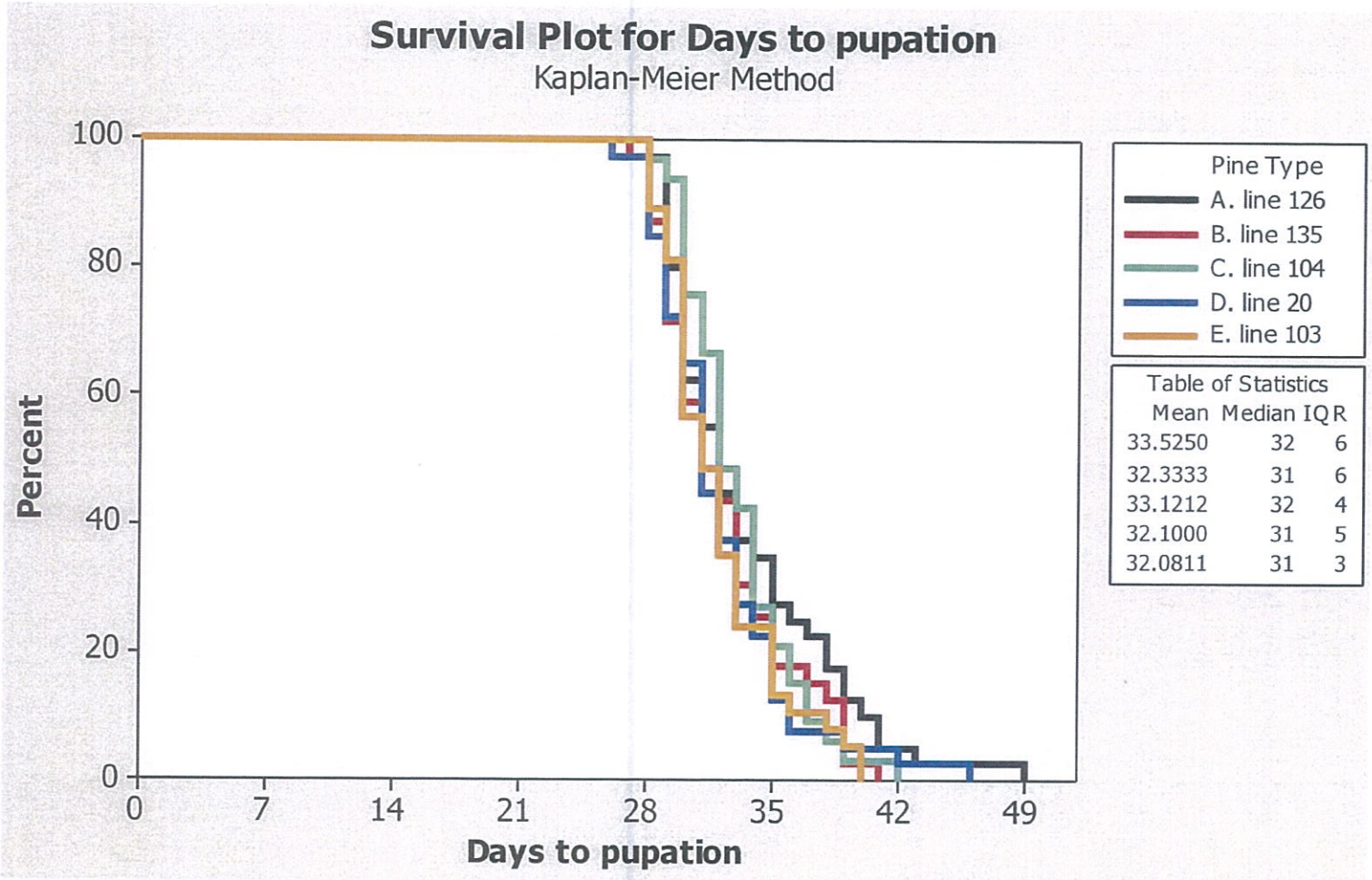


Figure 5. ***Pseudocoremia suavis*** **larval duration, i.e. days from egg hatch to pupation on different** ***Pinus radiata*** **lines. Data from the two replicates was pooled.**

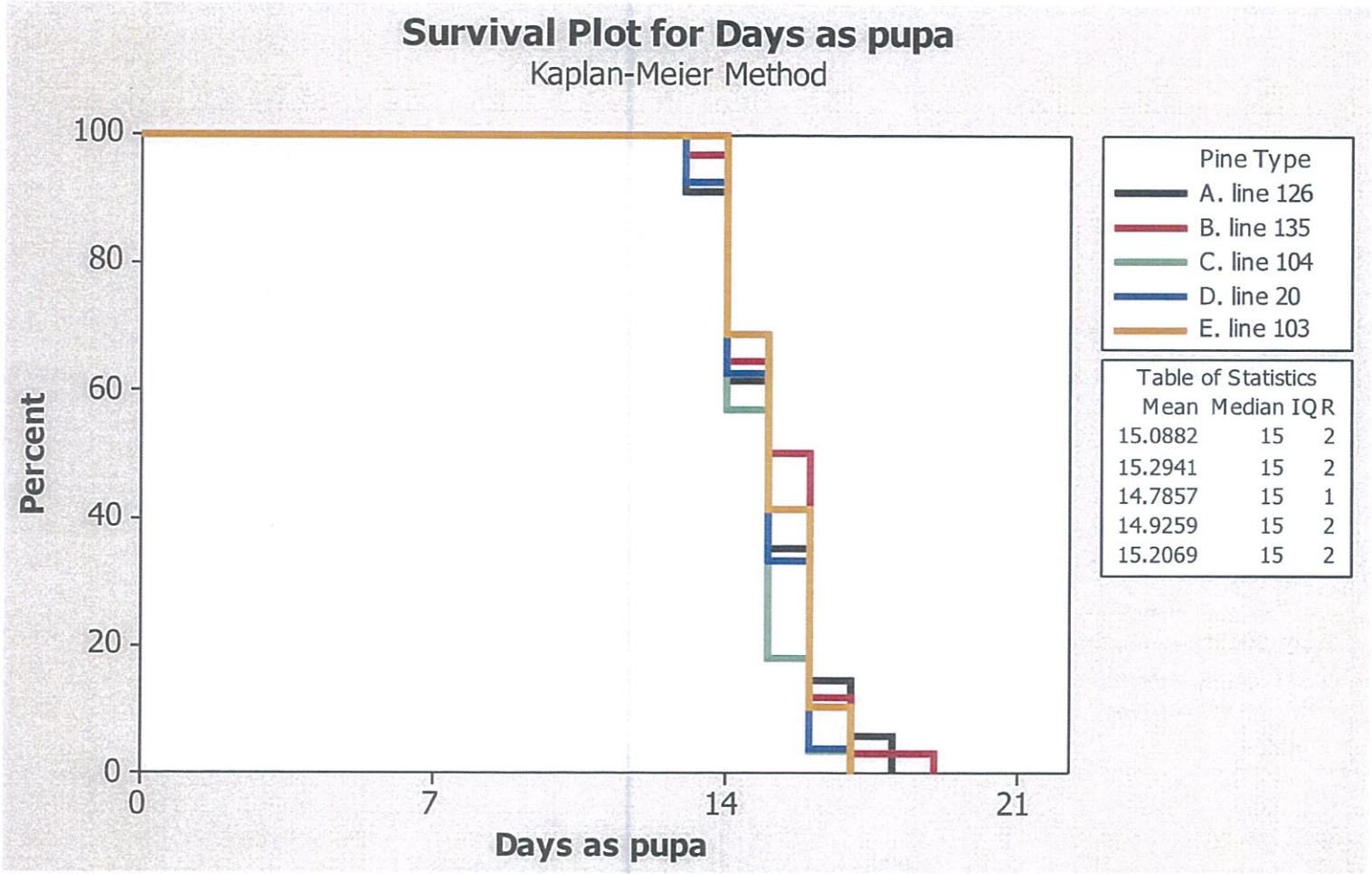


Figure 6. ***Pseudocoremia suavis*** **pupal duration, i.e. days from pupation to adult emergence on different** ***Pinus radiata*** **lines. Data from the two replicates were pooled.**

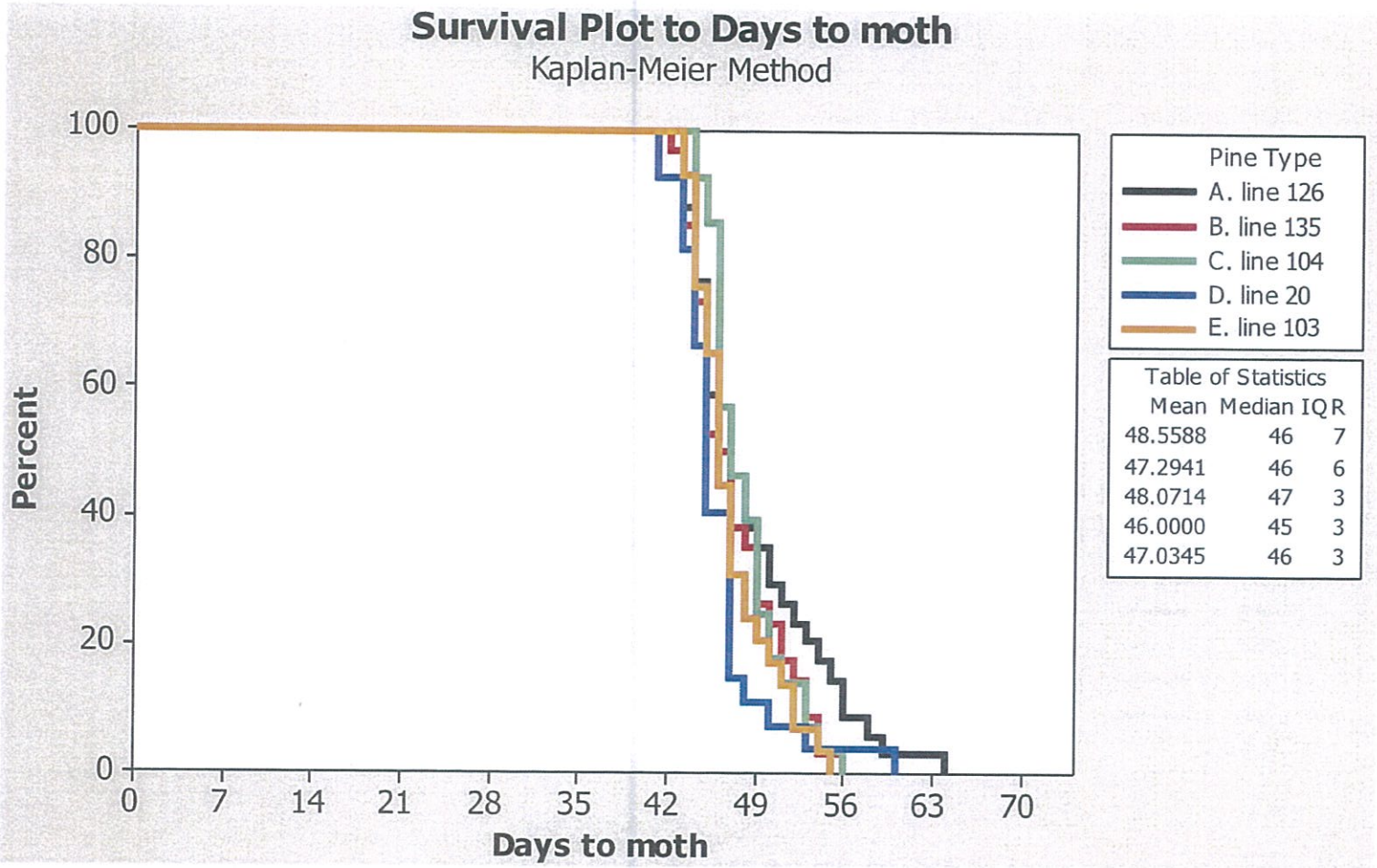


Figure 7. Days from egg hatch to adult emergence for *Pseudocoremia suavis* on different *Pinus radiata* lines. Data from the two replicates were pooled.

3.2.4 Pupal weight

In the analysis of pupal weights there was an interaction between pine line and replicate, meaning the effect of pine line on pupal weight was different between the two replicates. Separate ANOVA for the two replicates showed that most pine lines did not significantly differ in their effect, but, in replicate 1, the highest mean pupal weight (in Line 135) significantly differed from the lowest (in Line 104) ($F_{4,72} = 3.91$, $P = 0.006$). There were no significant differences in replicate 2 ($F_{4,101} = 2.09$, $P = 0.087$) (Figure 8). There was a replicate effect, replicate 2 having higher mean pupal weights in all pine lines, and a gender effect, with females having higher mean pupal weights in all pine lines.

Post hoc power analysis indicated that the power of the ANOVA to find a difference if a 10% difference in population means exists was somewhere between 35 and 50% (depending on statistical software used) in replicate 1, and in replicate 2 was 55-75%. This indicates that the difference between Lines 135 and 104 detected in replicate 1 did not reflect a consistent difference between the lines over time, or had become too small in replicate 2 to detect.

3.2.5 Absolute growth rate

The overall Absolute Growth Rate (AGR), defined here as pupal weight divided by days to pupate, was calculated because pupal weight and days to pupate on their own are only partial measures of an individual's success. Slower-growing larvae may be disadvantaged by a longer time to pupate but often reach a bigger pupal size.

The ANOVA showed a replicate effect, with replicate 2 yielding higher mean AGR in all pine lines. There was also an interaction between pine line and replicate, i.e. the effect of pine type on AGR was different between the two replicates (Figure 9). A separate ANOVA for replicate 1 showed a significant difference among the pine lines ($F_{4,72} = 2.96$, $P = 0.025$) but this difference between the highest and lowest was not significant when the Bonferroni multiple comparison procedure was used ($t = 2.885$, adjusted $P = 0.0516$). *Post hoc* power analysis indicated that a power of 30-50% to detect a difference if a 10% difference in population means existed.

In replicate 2, significant differences between pine lines were present ($F_{4,101} = 5.38$, $P = 0.007$), the mean AGR in Line 20 being higher than those on Lines 126 (isogenic control),104 and 135. The statistical power to detect differences was reasonably good in replicate 2, with a 70-85% chance of detecting a difference if a 10% difference in population means existed, suggesting that the differences detected may reflect genuine population differences not detected in the first replicate.

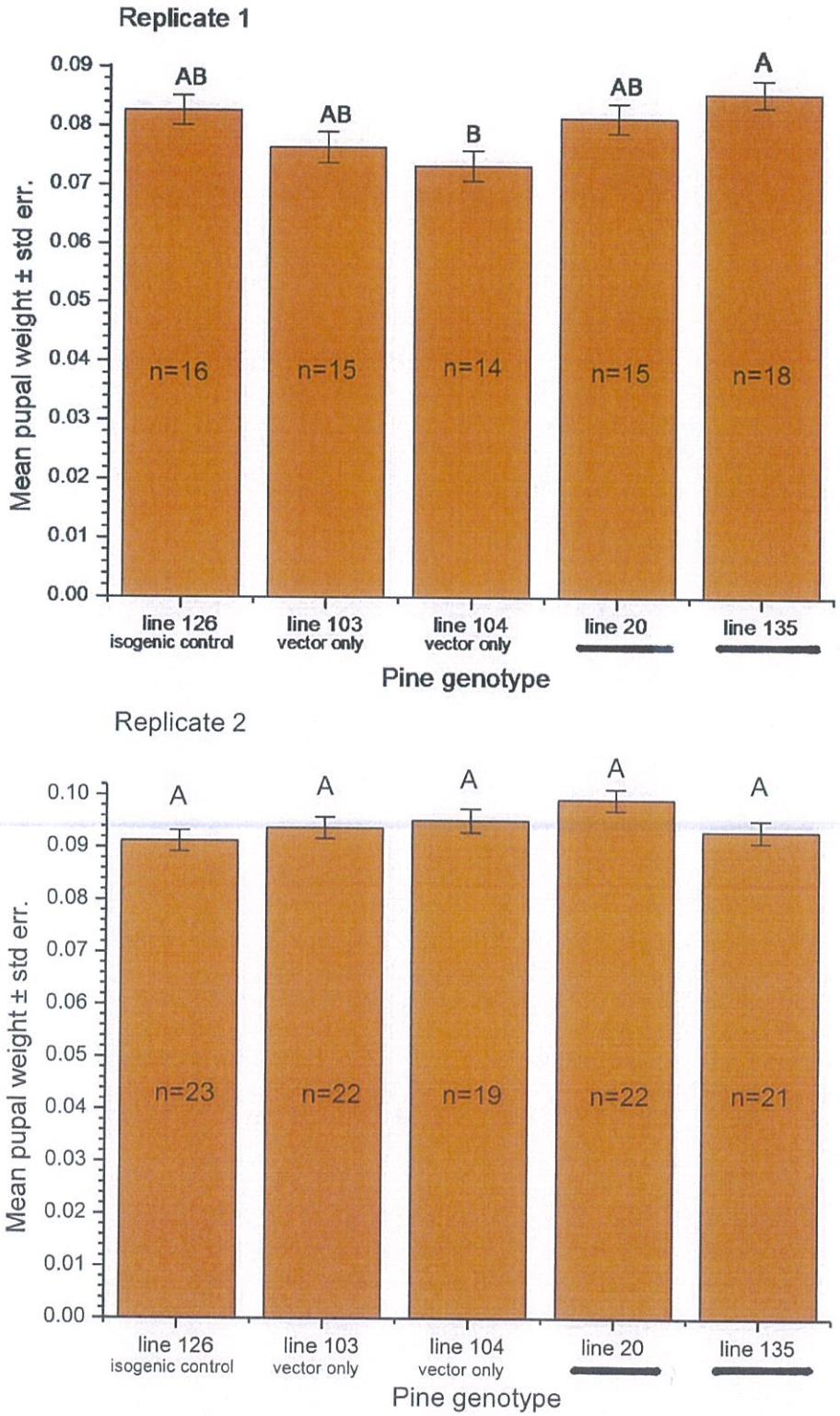


Figure 8. Mean pupal weights of *Pseudocoremia suavis* fed on five different *Pinus radiata* lines, plotted separately for replicates 1 and 2. Adjusted means from ANOVA are shown. Means sharing a letter are not significantly different (Bonferroni multiple comparison procedure at the 0.05 family level of significance).

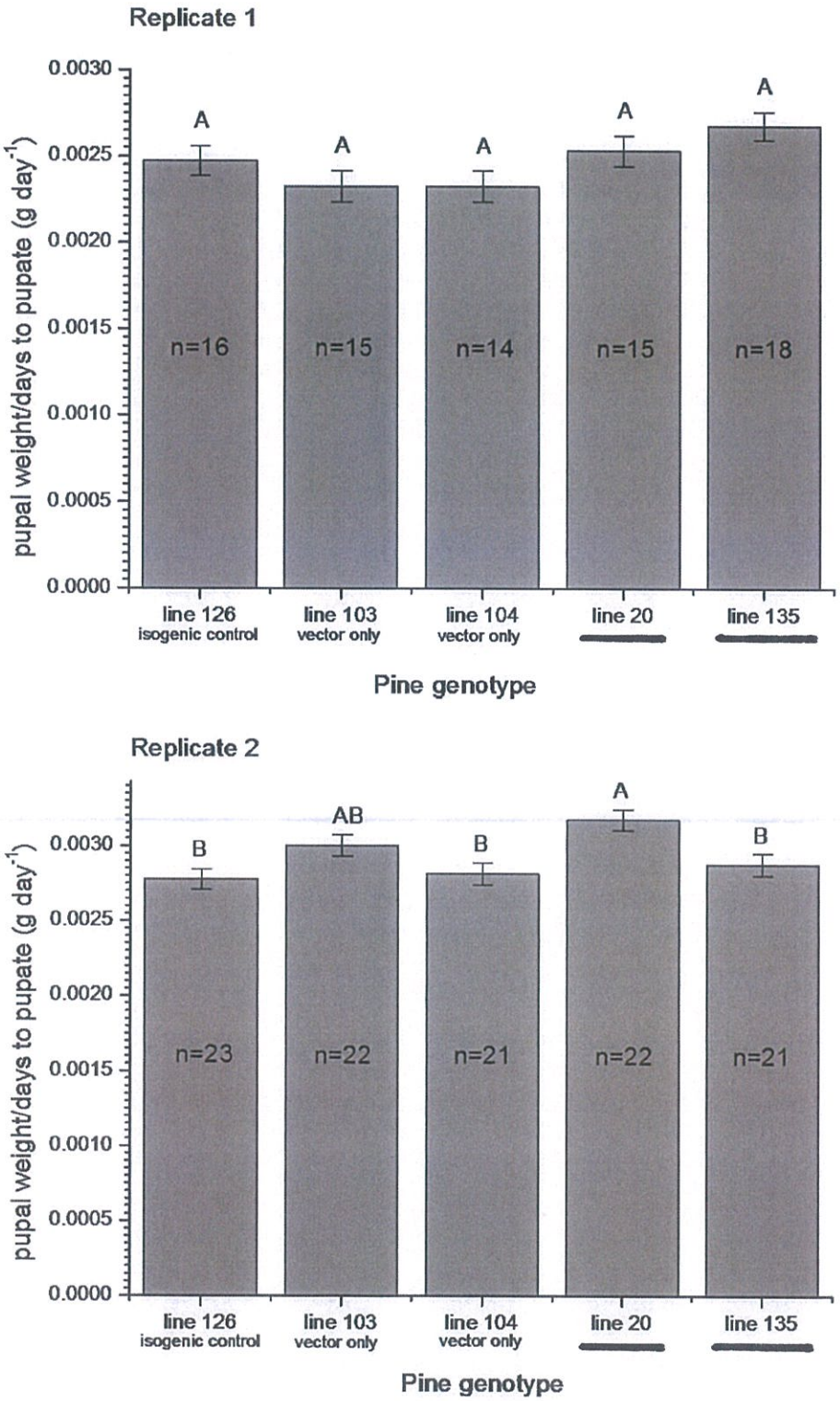


Figure 9. Absolute Growth Rate (Pupal weight / days to pupation) of *Pseudocoremia suavis* fed on five different *Pinus radiata* lines. Adjusted means from ANOVA are shown. Means sharing a letter are not significantly different (Bonferroni multiple comparison procedure at the 0.05 family level of significance).

4 RECOMMENDATION

We would recommend a third measurement of tree phenotype and a third bioassay replicate, dispensing with larval weigh-days in order to be able to increase numbers of larvae per treatment to 45. This increased number of larvae per treatment is possible now that the trees have grown to a size that would provide sufficient needle material to feed a larger number of larvae. Greater sample size could potentially provide adequate statistical power (at least 80%) to detect any meaningful differences in survival and development times, and give good or very good statistical power for comparisons of pupal weights and AGR. Small sample sizes are often adequate when treatments have large effects or only large effects are of interest. But when treatments are expected to have more subtle effects that are still of biological interest, as might be expected with insects feeding on pine larger sample sizes may be needed to detect these. In the case that there are no population differences, sample size is important for appropriately accepting the null hypothesis.