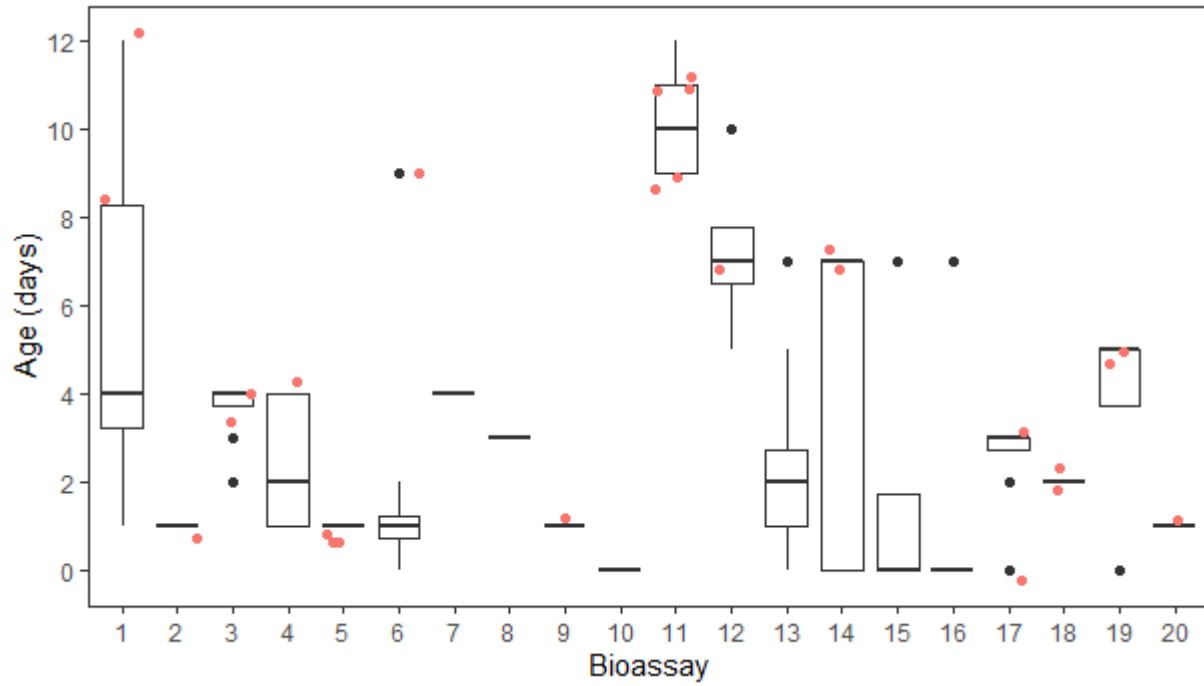
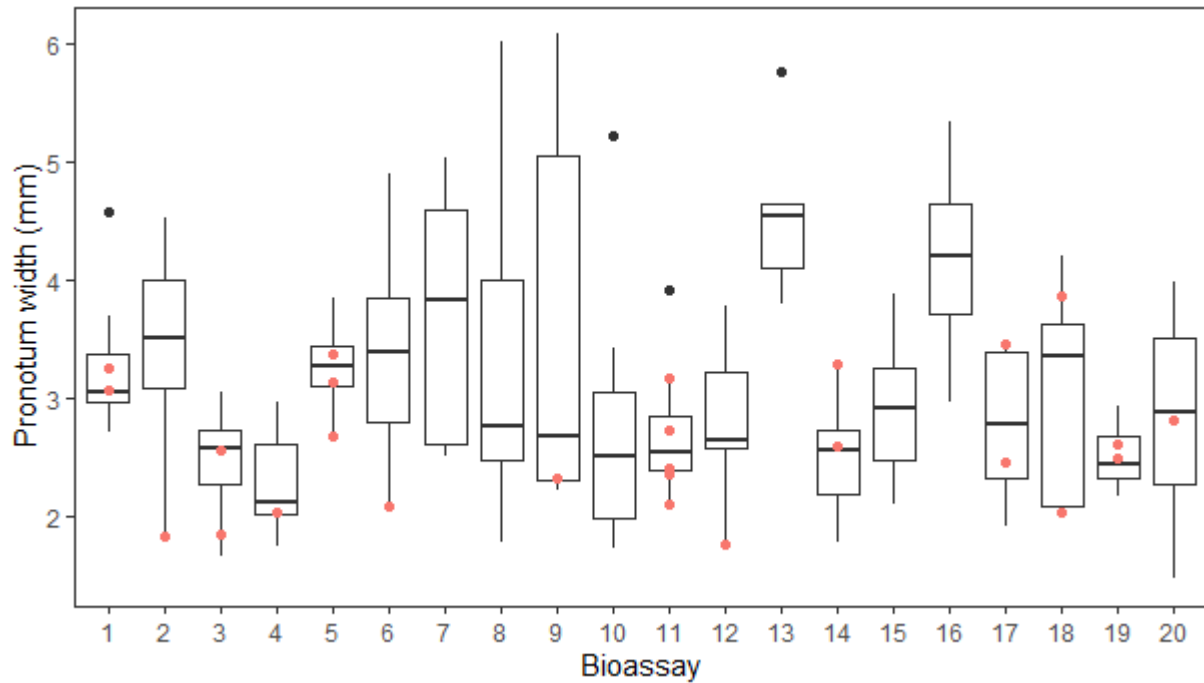


Supplementary materials



Supplementary Figure 1: Distribution of male age within each bioassay. Orange dots represent the age of the males that mated successfully. Points are scattered around values to avoid superposition.



Supplementary Figure 2: Distribution of male pronotum width within each bioassay. Orange dots represent the pronotum width of the males that mated successfully.

Model selection to model female attractiveness

Table 1 shows model set I. Using the AICc of each model we can rank the models from the model that best fits the data (smallest AICc) to the model that least fits the data (largest AICc) (**Table 1**, column 4). When using GLMMs, it is usually not recommended to use a hypothesis testing approach, but rather to use a model selection approach (Bolker et al 2009). As such, we determine a 95% confidence interval of models that best fit the data and use model averaging to obtain model estimates (Table 1, row 10). The first step is to calculate the Akaike weight (w_i) of each model (Table 1, column 6). This is done by dividing the relative likelihood of each model ($RL = \exp(-0.5 \times (AICc_i - AICc_{best\ model}))$) by the sum of the relative likelihoods of all models. We then gradually sum the Akaike weights of the models until the sum of Akaike weights > 0.95 . We start with the best model ($w_{16} = 0.546$) and add the Akaike weight of the second best model ($w_{16} + w_{19} = 0.729$), followed by the Akaike weight of the third best model and so on ($w_{16} + w_{19} + w_{12} + w_{18} = 0.968$). From this, we determine that models I6, I9, I2 and I8 all fall within the 95% confidence interval. The estimates of these four models can be averaged to determine model estimates. The averaged parameter estimate is the sum of the parameter estimate predicted by each model within the 95% confidence interval multiplied by the Akaike weight of the corresponding model. Another good indicator of how well the models within the 95% confidence interval fit the data compared to the rest of the tested models is the evidence ratio (Table 1, column 7). The evidence ratio of a model is the inverse of the relative likelihood and indicates how many times less likely the model is compared to the best model. Most notably, the evidence ratios showed that the null model (model I1) is 15 606 times less likely than the best model (model I6). This means that it is very unlikely that mating attempts are performed randomly by males, but rather that both male and female traits affect mate choice by males.

Model selection to model female receptivity

Table 2 shows model set J. Using the AICc of each model we can rank the models from the model that best fits the data (smallest AICc) to the model that least fits the data (largest AICc) (**Table 2**, column 4). We determined that all models except model J3 fall within the 95% confidence interval. The estimates of these eight models can be averaged to determine model estimates. Another good indicator of how well the models within the 95% confidence interval fit the data compared to the rest of the tested models is the evidence ratio (**Table 2**, column 7). The null model (model J1) is 10.61 times less likely than the best model (model J2) but is still within the 95% confidence interval. This means that while mating could be random, it is more likely that male and female trait affect mating success.