

# Impact of pedigree depth in the Australian strawberry breeding program: variance component estimation, prediction accuracies and estimation of clonal and breeding values

Dr. Ky Mathews PhD<sup>1\*</sup>, Dr. Katie O'Connor PhD<sup>1</sup> and Dr. Jodi Neal PhD<sup>1</sup>

<sup>1\*</sup>Department of Primary Industries, Queensland Government, Nambour, Australia, 4560, Queensland, Maroochy Research Facility, 47 Mayers Road.

\*Corresponding author(s). E-mail(s): [ky.mathews@daf.qld.gov.au](mailto:ky.mathews@daf.qld.gov.au);  
Contributing authors: [katie.oconnor@daf.qld.gov.au](mailto:katie.oconnor@daf.qld.gov.au);  
[Jodi.Neal@daf.qld.gov.au](mailto:Jodi.Neal@daf.qld.gov.au);

## Abstract

Key message: Less than 30 words summarising the main achievement.  
Abstract should be around 200 words.

**Keywords:** A matrix, relatedness, yield

## 1 Introduction

The Australian Strawberry Breeding Program (ASBP) is the nation's sole strawberry breeding initiative and is lead by the Queensland Department of Primary Industries. The program's objective is to develop public varieties that offer high profitability for producers paired with reduced costs of cultivation and satisfy consumer demands to drive repeat purchases, for Australia's three strawberry growing climatic regions: subtropical, temperate and Mediterranean. In the 2023/24 period, ASBP varieties accounted for about 90% of the total commercial subtropical strawberry plant sales across Australia, and 51% at the national level ([Hort Innovation, 2024](#)). Australia's

strawberry sector was valued at the farm gate at approximately \$507.5 million, producing around 80,741 tonnes of strawberries in the same timeframe (Hort Innovation, 2024).

Cultivated strawberry (*Fragaria x ananassa* Duchesne ex Rozier;  $2n = 8X = 56$ ) is an open-pollinated, highly heterozygous allo-octoploid, derived through hybridisation between *Fragaria virginiana* and *F. chiloensis* (Edger et al, 2019; Darrow, 1966). While early studies of inheritance in octoploid strawberry found meiotic behavior to be a mix between disomic and polysomic (e.g. Lerceteanu-Köhler et al, 2003), more recent studies have found that inheritance is disomic (Rousseau-Gueutin et al, 2008; Tennessen et al, 2014; Sargent et al, 2016; Hardigan et al, 2020) and that analyses employing diploid Mendelian genetics are appropriate (Pincot et al, 2021).

Yield and fruit weight are two of the most important traits in driving profitability in strawberry (Herrington et al, 2012). Increases in yield lead to an increase in gross margin (Herrington et al, 2012). In the subtropical region, increased early season (in April and May) production is a current aim of the ASBP due to the high market punnet prices compared to the rest of the season (Herrington et al, 2012). Large fruit sizes lead to substantial increases in gross margin, as the time to pick and pack a fruit does not change with size, but larger fruit will result in more punnets and thus be more profitable for growers compared to fewer punnets of small fruit (Herrington et al, 2012). Fruit weight tends to decrease as temperatures increase (Menzel, 2021), and so average fruit weight in August in the subtropical region and in December to February in the temperate region are particularly important factors in ASBP selection decisions. It is for these reasons that the ASBP evaluates total season yield and average fruit weight as well as for individual weeks and months.

The ASBP uses pedigree information in the routine analysis of all traits to estimate breeding values for parental selection. The total genetic effects for each genotype, the so-called “clonal values”, are also estimated from these models. The analyses are performed in the linear mixed model framework and follow Oakey et al (2006) and Oakey et al (2007) which uses the numerator relationship matrix generated from pedigree information.

*Theory? More levels of pedigree will lead to more accurate estimates and larger additive genetic variance? see Yang and Su (2016).*

There are a number of studies that have investigated the effect of various pedigree depths in breeding programs. Atkin et al (2009) investigated six different levels of sugar cane pedigree as well as varying numbers of years of trial data to determine the minimum levels of both to accurately estimate breeding values. Their evaluation methods included a comparison of Akaike Information Criterion (AIC) between models, additive genetic and additive-by-environment variance estimates, and correlations between breeding values of the full dataset and subsets. Their results found that a minimum pedigree depth of three generations and five years of trial data were required for their crop and dataset. In their simulation study of pig populations, Yang and Su (2016) compared estimates of genetic parameters and reliability of estimated breeding values (EBVs) for 1-24 generations included in their pedigree and with different numbers of sires and dams. In their simulations, three to five generations of phenotypic records plus one to three generations of ancestors produced EBVs as accurate as the

full dataset. In their development of the AGHmatrix R package, [Amadeu et al \(2016\)](#) also investigated pedigree depths of 5, 10, 20, and 30 years in blueberry. They found that shallow pedigrees underestimated actual kinship coefficients between individuals as well as incorrect determination of double reduction levels, a phenomenon that affects polyploidy organisms such as the autotetraploid blueberry, and thus estimation of genetic parameters ([Amadeu et al, 2016](#)).

Various papers analysing strawberry data have included relationship matrices constructed of somewhat arbitrary pedigree generation depths, ranging from four to 20 ([Whitaker et al, 2012](#); [Kennedy et al, 2014](#); [Osorio et al, 2014](#); [Paynter et al, 2014](#)). Currently, the breeding program uses a standard pedigree depth of 10 generations *out of a total of X generations*. It is of interest to determine what an optimal pedigree depth could be and the effect of changing pedigree depth on the genetic variance estimates, the breeding and clonal value estimates and the prediction accuracies these estimates.

The ASBP have implemented the contemporary group technology described in [Smith et al \(2021\)](#) to determine that for the subtropical program 5 years was sufficient, and for the temperate program 4 years (unpublished).

Repeated measures or longitudinal analysis of yield data does not appear to be standard practice in strawberry, with research groups instead summing individual harvests to obtain a total season yield (e.g. [Cockerton et al \(2021\)](#); [Gezan et al \(2017\)](#); [Fan et al \(2024\)](#); [Osorio et al \(2021\)](#)). Similarly for fruit weight, a common method is to divide total (marketable) yield by the total number of (marketable) fruit across a season (e.g. [Sleper et al \(2025\)](#); [Osorio et al \(2021\)](#); [Whitaker et al \(2012\)](#)), or average across all harvests to obtain a mean weight (e.g. [Prohaska et al \(2024\)](#)). Repeated measures analyses have been used in other horticultural crops to model traits measured temporally, for example in macadamia ([Hardner and De Faveri, 2023](#)), sweet cherry ([Piaskowski et al, 2018](#)) and... *Ky to insert her favourite examples (though perhaps we need to reword "horticultural")*.

The main purpose of this paper is to investigate the effect of depth of pedigree used to calculate the numerator relationship matrix on the additive and non-additive genetic effects, the breeding and clonal value estimates and their respective accuracies. Yield and fruit weight will be used in the study as example traits. A secondary purpose of this study is to investigate whether a repeated measures analysis is more appropriate in strawberries than analysing each harvest week individually and then summing or averaging data to obtain season values.

## 2 Materials and methods

### 2.1 Experimental design and phenotyping

### 2.2 Construction of A-matrices

Despite strawberry being octoploid, we used the option `ploidy = 2` when constructing the A matrix in AGHmatrix due to its diploid-behaving manner.

*If we do explore an octoploid A matrix and find that it is a better fit than diploid, then we should explore different levels of double reduction (the probability that two sister chromatids pass into the same gamete), as per [Amadeu et al \(2016\)](#).*

Recent studies have found that a small amount of genetic variance could be attributed to dominance ([Feldmann et al, 2024](#); [Sleper et al, 2025](#)). Thus, we also constructed a dominance relationship matrix in AGHmatrix using the option dominance = TRUE.

### 3 Results

Figures and tables are labeled with a prefix (fig or tab, respectively) plus the chunk label.

### 4 Discussion

The results of this work will be informative for a study regarding the level of missingness in the ASBP (unpublished.)

### 5 Conclusions

We found. . .

**Supplementary information.** If your article has accompanying supplementary file/s please state so here.

**Acknowledgments.** The authors thank present and past members of the Australian Strawberry Breeding Program and staff at the Maroochy Research Facility for the involvement in field preparation and maintenance, plant propagation, plant maintenance, fruit harvesting and processing, data collection, etc etc.

## Declarations

### 5.1 Funding

The Australian Strawberry Breeding Program has been funded by Hort Innovation using the strawberry research and development levy, with co-contributions from the Queensland Government through its Department of Primary Industries and funds from the Australian Government.

### 5.2 Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

### 5.3 Authors' Contributions

KM and KO developed the idea. KM wrote analytical R code. etc etc.

### 5.4 Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## References

- Amadeu RR, Cellon C, Olmstead JW, et al (2016) Aghmatrix: R package to construct relationship matrices for autotetraploid and diploid species: a blueberry example. *The plant genome* 9(3):plantgenome2016.01.0009. <https://doi.org/10.3835/plantgenome2016.01.0009>
- Atkin FC, Dieters MJ, Stringer JK (2009) Impact of depth of pedigree and inclusion of historical data on the estimation of additive variance and breeding values in a sugarcane breeding program. *Theoretical and Applied Genetics* 119(3):555–565. <https://doi.org/10.1007/s00122-009-1065-7>
- Cockerton HM, Karlström A, Johnson AW, et al (2021) Genomic informed breeding strategies for strawberry yield and fruit quality traits. *Frontiers in Plant Science* 12. <https://doi.org/10.3389/fpls.2021.724847>
- Darrow GM (1966) The strawberry history, breeding and physiology. The New England Institute for Medical Research

- Edger PP, Poorten TJ, VanBuren R, et al (2019) Origin and evolution of the octoploid strawberry genome. *Nature Genetics* 51(3):541–547. <https://doi.org/10.1038/s41588-019-0356-4>
- Fan Z, Verma S, Lee H, et al (2024) Strawberry soluble solids qtl with inverse effects on yield. *Horticulture Research* 11(2). <https://doi.org/10.1093/hr/uhad271>
- Feldmann MJ, Pincot DDA, Seymour DK, et al (2024) A dominance hypothesis argument for historical genetic gains and the fixation of heterosis in octoploid strawberry. *GENETICS* <https://doi.org/10.1093/genetics/iyae159>
- Gezan SA, Osorio LF, Verma S, et al (2017) An experimental validation of genomic selection in octoploid strawberry. *Horticulture Research* 4(Article 16070). <https://doi.org/10.1038/hortres.2016.70>
- Hardigan MA, Feldmann MJ, Lorant A, et al (2020) Genome synteny has been conserved among the octoploid progenitors of cultivated strawberry over millions of years of evolution. *Frontiers in Plant Science* 10. <https://doi.org/10.3389/fpls.2019.01789>
- Hardner C, De Faveri J (2023) Flexible linear mixed models for complex data in horticultural tree breeding. In: *Acta Horticulturae*, pp 139–146, <https://doi.org/10.17660/ActaHortic.2023.1362.19>
- Herrington ME, Wegener M, Hardner C, et al (2012) Influence of plant traits on production costs and profitability of strawberry in southeast queensland. *Agricultural Systems* 106(1):23–32. <https://doi.org/10.1016/j.agsy.2011.11.003>
- Hort Innovation (2024) Australian horticulture statistics handbook 2023/24
- Kennedy C, Osorio LF, Peres NA, et al (2014) Additive genetic effects for resistance to foliar powdery mildew in strawberry revealed through divergent selection. *Journal of the American Society for Horticultural Science* 139(3):310–316. <https://doi.org/10.21273/JASHS.139.3.310>
- Lerceteau-Köhler E, Guérin G, Laigret F, et al (2003) Characterization of mixed disomic and polysomic inheritance in the octoploid strawberry (*Fragaria* × *ananassa*) using AFLP mapping. *Theoretical and Applied Genetics* 107(4):619–628. <https://doi.org/10.1007/s00122-003-1300-6>
- Menzel C (2021) Higher temperatures decrease fruit size in strawberry growing in the subtropics. *Horticulturae* 7(2):34. <https://doi.org/10.3390/horticulturae7020034>
- Oakey H, Verbyla A, Pitchford W, et al (2006) Joint modeling of additive and non-additive genetic line effects in single field trials. *Theor Appl Genet* 113:809–819

- Oakey H, Verbyla AP, Cullis BR, et al (2007) Joint modeling of additive and non-additive (genetic line) effects in multi-environment trials. *Theor Appl Genet* 114(8):1319–32. <https://doi.org/10.1007/s00122-007-0515-3>
- Osorio LF, Pattison JA, Peres NA, et al (2014) Genetic variation and gains in resistance of strawberry to *Colletotrichum gloeosporioides*. *Phytopathology* 104(1):67–74. <https://doi.org/10.1094/phyto-02-13-0032-r>
- Osorio LF, Gezan SA, Verma S, et al (2021) Independent validation of genomic prediction in strawberry over multiple cycles. *Frontiers in Genetics* 11(1862). <https://doi.org/10.3389/fgene.2020.596258>
- Paynter ML, De Faveri J, Herrington ME (2014) Resistance to *Fusarium oxysporum f. sp. fragariae* and predicted breeding values in strawberry. *Journal of the American Society for Horticultural Science* 139(2):178–184. <https://doi.org/10.21273/JASHS.139.2.178>
- Piaskowski J, Hardner C, Cai L, et al (2018) Genomic heritability estimates in sweet cherry reveal non-additive genetic variance is relevant for industry-prioritized traits. *BMC Genetics* 19(1):23–38. <https://doi.org/10.1186/s12863-018-0609-8>
- Pincot DD, Ledda M, Feldmann MJ, et al (2021) Social network analysis of the genealogy of strawberry: retracing the wild roots of heirloom and modern cultivars. *G3* 11(3):jkab015. <https://doi.org/10.1093/g3journal/jkab015>
- Prohaska A, Rey-Serra P, Petit J, et al (2024) Exploration of a European-centered strawberry diversity panel provides markers and candidate genes for the control of fruit quality traits. *Horticulture Research* 11(7). <https://doi.org/10.1093/hr/uhae137>
- Rousseau-Gueutin M, Lerceteau-Köhler E, Barrot L, et al (2008) Comparative genetic mapping between octoploid and diploid fragaria species reveals a high level of colinearity between their genomes and the essentially disomic behavior of the cultivated octoploid strawberry. *Genetics* 179(4):2045–2060. <https://doi.org/10.1534/genetics.107.083840>, URL <http://dx.doi.org/10.1534/genetics.107.083840>
- Sargent D, Yang Y, Šurbanovski N, et al (2016) Haplo-snp affinities and linkage map positions illuminate subgenome composition in the octoploid, cultivated strawberry (*Fragaria* × *ananassa*). *Plant Science* 242:140–150. <https://doi.org/10.1016/j.plantsci.2015.07.004>
- Sleper J, Tapia R, Lee S, et al (2025) Within-family genomic selection in strawberry: Optimization of marker density, trial design, and training set composition. *The Plant Genome* 18(1). <https://doi.org/10.1002/tpg2.20550>
- Smith A, Ganesalingam A, Lisle C, et al (2021) Use of contemporary groups in the construction of multi-environment trial datasets for selection in plant breeding

- programs. *Frontiers in Plant Science* 11:623586. <https://doi.org/10.3389/fpls.2020.623586>
- Tennessen JA, Govindarajulu R, Ashman TL, et al (2014) Evolutionary origins and dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps. *Genome Biology and Evolution* 6(12):3295–3313. <https://doi.org/10.1093/gbe/evu261>
- Whitaker VM, Osorio LF, Hasing T, et al (2012) Estimation of genetic parameters for 12 fruit and vegetative traits in the university of florida strawberry breeding population. *Journal of the American Society for Horticultural Science* 137(5):316–324. <https://doi.org/10.21273/jashs.137.5.316>
- Yang H, Su G (2016) Impact of phenotypic information of previous generations and depth of pedigree on estimates of genetic parameters and breeding values. *Livestock Science* 187:61–67. <https://doi.org/10.1016/j.livsci.2016.03.001>