



# Genetic analysis of multiple fruit quality traits in mango across sites and years

Ian S. E. Bally · Joanne De Faveri

Received: 8 August 2020 / Accepted: 9 December 2020  
© Crown 2021

**Abstract** A key strategy to improve fruit quality and consumer appeal in mangoes is to breed trait improvements into new cultivars. There are several key fruit quality traits in mango. Knowledge of the heritability and relationship among these traits is important for breeding success. This paper implements a linear mixed model analysis including paternal information for analysis of 13 important fruit quality traits from mango cultivars planted across three environments and assessed over several seasons. The traits analysed were average fruit weight, skin background colour, blush colour, percentage blush, blush intensity, skin thickness, beak shape, stem-end shape, deformities, mesocarp colour, mesocarp texture, seed width and mesocarp recovery. The analysis allows investigation into trait heritabilities and stability of traits over years and sites, as well as genetic correlations among traits.

**Keywords** BLUP · Heritability · Fruit quality traits · Genetic correlation

## Introduction

The global production of mangoes is approximately 43.3 Million metric tons produced from 105 countries (Galan Saco 2017). Australian mango production is less than 0.2% of global production with approximately 61,474 tonnes produced annually, with a gross value of production (GVP) of \$195 million. Mangoes are grown across the tropical and subtropical regions of Queensland, the Northern Territory and Western Australia. 89% of all Australian production is sold and consumed domestically, with 83% sold as fresh fruit and 6% as processed (Horticulture Innovation Australia 2018).

Genetic improvement of crops through breeding is a key strategy for delivery of sustainable improvement in production efficiency and product quality. Many improved cultivars arise from breeding programs in India, USA, Israel, Brazil, Australia and South Africa (Bally and Dillon 2018; Iyer and Schnell 2009). Most of these programs aim to improve tree productivity, tree architecture and fruit quality such as fruit size, colour and flavour (Iyer and Schnell 2009; Bally and Dillon 2018). The ease of vegetative reproduction in mango allows the efficient capture and exploitation of genetic gain at any stage of a hybridisation program.

In Australia several factors have been identified as limiting mango industry growth, including the appeal of mangoes to consumers (fruit colour, flavour, aroma and mesocarp texture), seasonality of fruit, limited

---

I. S. E. Bally (✉) · J. De Faveri  
Queensland Department of Agriculture and Fisheries,  
PO Box 1054, Mareeba, QLD 4880, Australia  
e-mail: Ian.bally@daf.qld.gov.au

productive capacity of some cultivars, low first grade pack-out and access to new cultivars and orchard systems (Horticulture Innovation Australia 2017; AMIA 2014).

The breeding objectives of the Australian Mango Breeding Program are aligned with industry needs and include improvement of tree productivity, architecture, disease resistance and fruit quality traits of fruit size, colour and flavour (Kulkarni et al. 2002; Bally et al. 2017, 2013; Bally 2008).

To achieve these objectives the Australian Mango Breeding Program is combining genetic traits of different cultivars through controlled hand pollination techniques to generate new hybrid progeny. Accurate parental identification and evaluation of progeny's phenotypic performance across a range of environments has enabled analyses of multiple fruit traits for their heritabilities, breeding values and correlation among traits (Bally et al. 2009b; Hardner et al. 2012). These genetic relationships are useful to guide the selection of suitable next generation parents from the families represented in the breeding program.

The optimal selection of future parents to maximize progress in desirable traits is a key factor in any breeding program. Evaluating the breeding value or additive genetic effect for a cultivar for a given trait gives the expected average performance of progeny derived from crosses using this cultivar as a parent (Falconer and Mackay 1996). By selecting potential parents based on their breeding values for a key trait, the resulting progeny are likely to have improved values of this trait.

The heritability of a trait represents the proportion of variation in the phenotype that is due to genetic factors, with the narrow sense heritability being the proportion of phenotypic variance that can be attributed to additive genetic variance. The narrow sense heritability is important in plant breeding as it determines the amount of progress that can be made by selecting and crossing the best individuals in a population (Bernardo 2010). Traits with higher narrow sense heritability are likely to provide greater response to selection.

This paper presents a statistical genetic analysis of multi-site, multi-year data from multiple key traits from the Australian Mango Breeding Program. The analysis approach is based on linear mixed models including pedigree information and factor analytic models (Smith et al. 2001) for determining the genetic

covariance structure over sites and years. The analysis provides predictions of breeding values (BLUPs) and narrow sense heritabilities for each trait and allows investigation into genotype by environment interaction. The analysis also provides insight into the genetic correlation among traits. A similar mixed model approach was implemented in the univariate analysis of mango fruit weight (Hardner et al. 2012).

## Methods

### Genetic material and trial environments

The trees evaluated in this study consisted of 1719 hybrids (progeny) from 39 families, generated by crossing 29 parents in a sparse design (Hardner et al. 2012). Hybrids were generated using hand pollination techniques (Bally et al. 2009a). Hybrid progeny were planted at three sites across Northern Australia, at Coastal Plains Horticultural Research Farm, Darwin in the Northern Territory (NT), at Southedge Research Station, Mareeba, in Queensland (QLD) and at Frank Wise Institute, Kununurra in Western Australia (WA). Hybrid seedlings were planted in the NT while budwood from these seedlings was grafted on to Kensington Pride rootstock and planted in QLD and WA. Families with between one to 138 progeny were analysed (Table 1).

Subsets of hybrid progeny were assessed at the 3 locations over 6 years (QLD: 1999 to 2005; NT: 2000 to 2002 and 2004; WA: 2000 to 2005). Individual progeny were assessed, in at least 2 of the 6 years, resulting in an unbalanced sampling method (Table 2).

### Fruit quality traits

Thirteen fruit quality traits were assessed on each hybrid progeny, including five traits with continuous rating scales and eight traits with ordered categorical rating scales with four or more categories (Table 3). Each trait is described below.

### Average fruit weight

The average fruit weight per tree was calculated from five fruit harvested from each tree at full maturity and weighed at the eating ripe stage. Average fruit weight data were analysed as continuous quantitative data.

**Table 1** Parents and the number of progeny per family assessed across the 3 locations (Mareeba, Darwin and Kununurra). Numbers in brackets after parent names are the numerical identification of the parents used in Fig. 2

Female parent	Male parent						
	Banana callow (3)	Hybrid 10 (8)	Kensington pride (16)	Nam doc mai (19)	R2E2 (22)	Sensation (23)	Tommy Atkins (25)
Alphonso (1)					110		
Apple (2)		138	124				
Creeper (4)			1				
Edward (5)			3				
Early Gold (6)		33	51				
Glenn (7)	9		51				
Hybrid 17 (9)			83				1
Haden (10)		50	117	83	4		2
Irwin (11)			3				
Julie (12)			102				
Keitt (13)			44			3	
Kent (14)					2	1	
Kensington Mono (15)			76				
Kensington Pride (16)			14				
Lippens (17)			1				
Magovar (18)						1	
Nam Doc Mai (19)		68	67				
Padiri (20)			21				
Palmer (21)		12	58				
Suvarnareka (24)			100				
Tommy Atkins (25)			55				
Van Dyke (26)			2				
Vellai Columban (27)			78		115		
Willard (28)	3		3				
Zill (29)			41				

**Table 2** The number of hybrid progeny assessed each year at each location used in the analyses

Assessment year	Queensland	Northern Territory	Western Australia	Total
1999–00	356	0	0	356
2000–01	417	199	82	698
2001–02	590	243	112	945
2002–03	445	562	94	1101
2003–04	322	0	108	430
2004–05	470	637	170	1277
Total	2600	1641	566	4807

**Table 3** Thirteen fruit quality traits that were analysed, rating scales and rating levels. Five traits with continuous rating scales and eight traits with ordered categorical rating scales

Description	Scale	Unit/Rating levels
Average fruit weight	Continuous	grams
Skin background colour	Categorical	4
Blush colour	Categorical	4
Percentage blush	Continuous	%
Blush intensity	Categorical	5
Skin thickness	Continuous	mm
Beak shape	Categorical	5
Stem-end shape	Categorical	5
Deformities	Categorical	4
Mesocarp colour	Categorical	5
Mesocarp texture	Categorical	5
Seed width	Continuous	mm
Mesocarp recovery	Continuous	mm

#### Skin background colour

The skin ground colour is the underlying green/ yellow/ orange colour of the fruit skin at eating ripe. This colour does not include the blush colour of the fruit. The skin ground colour transitions from green to yellow as the fruit ripen and chlorophyll is lost from the fruit skin (Medlicott et al. 1986). Skin ground colour was categorised according to the predominant colour of the un-blushed skin at eating ripe and analysed in order of most to least desirable as either yellow, orange, green/yellow, or green. Some cultivars do not de-green fully during ripening resulting in an undesirable blotchy green, yellow appearance.

#### Blush colour

Mango fruit blush colours range from orange, through pink, to red and purple from anthocyanin pigments that result from the activation of cyanidin-O-galactoside synthesis stimulated by direct exposure to sunlight (Berardini et al. 2005a). Mango fruit blush colour was rated on an ordered categorical scale, in order of most to least desirable, as burgundy, red, pink or orange.

#### Percent blush coverage

The percentage of fruit skin covered with blush was assessed by visually estimating the percentage of blush separately on both sides of the fruit at eating ripe stage and taking the average. The percentage blush was analysed as continuous quantitative data with higher percentages preferable.

#### Bush intensity

The blush colour on the fruit skin can vary not only by the amount of skin covered, but also by the intensity of the blush colour. The more intense the blush colour the more it completely covers the underlying ground colour of the fruit. Blush intensity was scored as an ordered categorical rating in order of most to least desirable as: medium intensity similar to the cultivar 'Haden', slight intensity, similar to the cultivar 'Kensington Pride', solid intensity, similar to the cultivar 'Tommy Atkins', barely visible, or no blush. Blush intensity data from Western Australia was not included in the multi-trait analyses.

#### Skin thickness

The thickness of mango fruit skin influences the total mesocarp (flesh) recovery of the fruit and how easy it is to peel the skin from the fruit. Skin thickness was measured in mm at the eating ripe stage using a digital calliper. Skin thickness was calculated as the average skin thickness of five measurements taken randomly around the longitudinal circumference after removing the fruit cheek from the seed on each of the five fruit in the sample.

#### Beak shape

The shape of mango fruit vary from round to elongate and with fruit colour are the most recognisable features of a mango cultivar for consumers. The beak shape describes the amount the styler end of the fruit protrudes and is a significant component of fruit shape. Beak shape was scored in ripe fruit, in an ordered categorical rating scale, in order of most to least desirable, as absent, very slight, slight, medium, or prominent. Beak shape data from Western Australia was not included in the multi-trait analyses.

### Stem-end shape

The stem-end shape of a mango fruit influences the shape of the fruit and the amount of detritus material and moisture that accumulates externally at the stem end of the fruit during growth. Depressed stem-ends accumulate more material that can blemish the fruit and cause a degrading of fruit quality. Stem-end shape was scored an ordered categorical rating scale, in order of most to least desirable, as level, slightly depressed, slightly raised or highly depressed. Stem-end shape data from Western Australia was not included in the multi-trait analyses.

### Deformities

Fruit deformities appear as lumps on the fruit or as misshapen fruit which are unmarketable. Fruit deformities were rated on an ordered categorical scale, with lower valued preferred as either, none, slight, medium or many.

### Mesocarp colour

The colour of the fruit mesocarp (flesh) ranges from pale yellow green to dark orange in ripe fruit. Both carotenoids and anthocyanins contribute to the intensity of the mesocarp colour (Proctor and Creasy 1969; Pott et al. 2003). Mesocarp colour was scored, in order of most to least desirable, on a one to five ordered categorical scale using colour patch cards (The Royal Horticultural Society 2001) as either orange group 24A, yellow orange group 32A, yellow group 15A, yellow group 13 B, or yellow group 6A.

### Mesocarp texture

Mesocarp texture refers to the firmness and fibre associated with the fruit mesocarp. Firm, low fibre textures are preferable to soft fibrous textures. Mesocarp texture was scored, in order of most to least desirable, using an ordered categorical rating scale based on commonly known cultivars as either soft or no fibre (c.v. 'Nam Doc Mai'), soft and low fibre (c.v. 'Kensington Pride') firm and medium fibre (c.v. 'R2E2'), firm and stringy (c.v. 'Tommy Atkins') and soft and stringy (c.v. 'Common').

### Seed width

Seed width refers to the width of the seed and leathery endocarp, often referred to as the stone. Seeds (embryos enclosed in their leathery endocarp) were extracted from the ripe fruit samples, measured with digital callipers in mm and analysed as continuous quantitative data. Thinner seeds are seen as more desirable as they increase mesocarp recovery.

### Mesocarp recovery

The mesocarp (flesh) recovery refers to the percentage of edible mesocarp that can be extracted from the fruit. Higher percentages of mesocarp recovery are preferred. Mesocarp recovery was calculated by subtracting the seed and skin weight from the fruit weight and expressing it as a percentage of the fruit weight as follows:

$$\text{Flesh recovery} = \frac{(\text{Fruitweight} - (\text{Seed} + \text{Skinweight}))}{\text{Fruitweight}} \times 100$$

### Statistical methods

Prior to analysis, each of the categorical variables were transformed to a numerical rating scale with higher ratings associated with more desirable fruit. Then all traits were analysed individually across sites and years using a multi-environment (MET) multi-harvest analysis (Hardner et al. 2012; De Faveri 2013) using linear mixed models incorporating factor analytic models (Smith et al. 2001; Meyer 2007) and pedigree information. The models were fitted using ASReml-R (Butler et al. 2009). Residuals were investigated and assessed to meet the assumptions for analysis.

The aim of the linear mixed model analysis was to predict additive genetic effects (BLUPs) for each cultivar for each trait and genetic and residual variance components for estimation of genetic correlations and narrow sense heritabilities for each trait across harvest seasons within trials. The linear mixed model used for analysis of each trait was of the form:

$$y = X\tau + Z_o u_o + Z_g u_g + Z_f u_f + e$$

where  $y$  is the vector of observations, fixed effects are given by  $X\tau$ , random (non-genetic) effects are given

by  $Z_o u_o$ , the random additive genetic effects by  $Z_g u_g$ , random family effects by  $Z_f u_f$  and the residual effects by  $e$ . It is assumed that  $e$  is normally distributed with zero mean and covariance matrix  $R$ .

The additive genetic effects  $u_g$  are assumed to be normally distributed and have mean zero and are independent of other random effects. The multi-site/multi-year model used in this paper treats the site by year combinations as a single component. It is assumed that the variance matrix of  $u_g$  is given by:

$$\text{var}(u_g) = G_s \otimes A$$

where  $G_s$  is the genetic covariance matrix for the site by year combinations,  $\otimes$  is the Kronecker product, and  $A$  is the additive relationship matrix as determined by the pedigree.

The random additive genetic effects were correlated across sites and years and BLUPs calculated for each cultivar across sites and years. The genetic covariance matrix  $G_s$  consisting of genetic variances for each site by year and genetic covariances between site by year combinations was modelled using factor analytic models (Smith et al. 2001). The factor analytic model provides a parsimonious approximation to the fully unstructured covariance model (Kelly et al. 2007). The order of factor analytic model required for each trait was determined using REML likelihood ratio tests (REMLRT).

In multi-environment trials the full residual covariance matrix  $R$  is typically given by a block diagonal matrix where  $R_j$  is the residual variance matrix for the  $j^{\text{th}}$  trial:

$$R = \text{diag}(R_j)$$

Therefore, each trial has its own residual covariance structure and residuals are assumed independent among trials. In this study the residual structure for each trial has been modelled using a diagonal variance matrix for each site, giving a separate residual variance for each year. Spatial analyses were not performed as at any one time only a selection of non-contiguous trees were measured. Models fitting more structured temporal residual correlation structures were investigated but were not significant or unable to be fitted most likely because of insufficient individual trees being measured at consecutive times.

The genetic covariance matrix ( $G_s$ ) (giving the genetic variances and correlations for the 15 site by

year combinations) was estimated for each trait to investigate the stability of traits across sites and years. Heat maps were constructed to visualise these genetic correlations and to assist in interpretation of the covariance matrices from the factor analytic models (De Faveri et al. 2015; Cullis et al. 2010).

Variance components from the model were used to estimate the narrow sense heritability for each trait for each year by trial combination using the following formula:

$$\hat{h}^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_f^2 + \hat{\sigma}_e^2}$$

where  $\hat{\sigma}_a^2$  was the estimated additive genetic variance,  $\hat{\sigma}_f^2$  was the estimated family variance and  $\hat{\sigma}_e^2$  was the estimated residual variance for the trait at a particular year within a trial.

Breeding values (random additive genetic effects) were predicted for each line for all years and trials for each trait separately. A principal component analysis was performed and biplot constructed on the trait by cultivar BLUPs (predicted over sites and years) to investigate the relationships among multiple traits and represented in a biplot (Fig. 4), generated using the statistical package R (R Core Team 2015).

## Results

The mean and standard errors based on the raw data for each trait by harvest within years and sites are presented in Table 4. It can be seen that some trait means differed across sites, for example average fruit weight was consistently higher in QLD than WA or NT. WA trials showed higher variation in average fruit weight between seasons than the other sites (Table 4).

The analyses of each trait were based on factor analytic models for modelling the genetic effects over sites and years. The order of factor analytic model (number of factors) was determined using REML likelihood ratio tests. For most traits a model with two factors (FA2) was deemed best, while for percent blush and average fruit weight a model with three factors (FA3) was chosen as the best model and for mesocarp recovery a model with four factors (FA4) was used.

**Table 4** Means and standard errors (se) for site by harvest year (rows) for each trait (columns)

Environment	Average fruit weight	se	Mesocarp colour	se	Blush percentage	se	Mesocarp recovery	se	Stem-end shape	se	Deformities	se
NT 2000–01	383.39	10.84	2.9	0.06	23.27	1.59	57.34	0.82	4.34	0.07	3.93	0.02
NT 2001–02	269.44	9.67	2.65	0.05	12.76	1.08	52.26	0.74	3.79	0.07	3.5	0.05
NT 2002–03	357.66	6.11	2.89	0.03	19.28	0.89	54.99	0.42	3.95	0.04	3.41	0.03
NT 2004–05	361.35	5.74	3.02	0.03	24.4	0.79	56.28	0.4	4.03	0.04	3.37	0.03
QLD 1999–00	431.97	9.92	2.63	0.04	36.75	1.22	60.16	0.58	4.04	0.05	3.23	0.04
QLD 2000–01	410.25	8.75	2.66	0.04	32.14	1.11	59.57	0.57	3.84	0.05	3.52	0.03
QLD 2001–02	438.76	7.63	2.63	0.04	24.65	0.87	62.06	0.49	3.74	0.05	3.13	0.03
QLD 2002–03	423.26	8.78	2.47	0.04	25.64	0.88	59.12	0.52	3.75	0.06	3.02	0.03
QLD 2003–04	425.16	8.87	2.73	0.05	22.54	1.21	62.15	0.6	4.01	0.06	2.66	0.05
QLD 2004–05	419.08	7.17	2.91	0.03	27.91	0.66	60.31	0.48	3.81	0.05	3.22	0.03
WA 2000–01	242.99	14.09	2.37	0.09	23.32	1.77	46.78	1.22	3.65	0.13	3.06	0.09
WA 2001–02	361.06	13.74	2.72	0.07	10.71	1	56.23	1.15	4.12	0.09	3.24	0.07
WA 2002–03	380.97	14.64	2.87	0.08	6.6	1.07	56.04	1.32	4.21	0.09	3.44	0.06
WA 2003–04	410.9	14	2.69	0.07	4.86	0.85	59.73	1.16	4.19	0.07	3.59	0.05
WA 2004–05	361.8	9.86	3.22	0.05	8.5	0.87	56.16	0.84	4.08	0.08	3.68	0.05
Environment	Beak shape	se	Skin blush intensity	se	Mesocarp texture	se	Skin background colour	se	Skin thickness	se	Seed width	se
NT 2000–01	3.6	0.1	2.61	0.09	3.21	0.07	3.6	0.04	0.9	0.03	19.09	0.2
NT 2001–02	3.35	0.08	2.29	0.08	3.31	0.08	3.51	0.05	1.27	0.03	19.15	0.19
NT 2002–03	3.88	0.05	2.29	0.05	3.47	0.05	3.55	0.03	1.01	0.02	18.97	0.11
NT 2004–05	3.97	0.04	2.83	0.04	3.67	0.05	2.89	0.05	0.93	0.01	18.61	0.1
QLD 1999–00	3.59	0.06	2.63	0.05	3.99	0.05	3.07	0.06	0.77	0.01	19.25	0.14
QLD 2000–01	3.86	0.05	2.84	0.05	3.8	0.05	3.31	0.04	0.85	0.01	19.43	0.13
QLD 2001–02	3.87	0.04	2.32	0.05	3.89	0.05	3.07	0.04	0.86	0.01	19.24	0.13
QLD 2002–03	3.52	0.06	2.23	0.05	3.13	0.07	3.37	0.04	0.71	0.01	19.08	0.14
QLD 2003–04	3.6	0.07	2.48	0.07	3.33	0.08	3.07	0.06	0.81	0.02	19.46	0.15
QLD 2004–05	3.82	0.05	2.56	0.05	3.61	0.06	3.08	0.05	0.91	0.01	19.19	0.12
WA 2000–01	3.18	0.15	2.98	0.12	3.88	0.09	3.46	0.07	1.26	0.04	19.39	0.25
WA 2001–02	3.82	0.12	2.45	0.11	3.67	0.12	2.89	0.09	0.6	0.03	21	0.26
WA 2002–03	3.77	0.12	2.13	0.13	3.66	0.13	3.19	0.08	0.46	0.02	20.57	0.26
WA 2003–04	3.82	0.11	2.32	0.12	4.17	0.09	3	0.09	0.46	0.02	19.96	0.25
WA 2004–05	3.82	0.09	2.15	0.1	3.75	0.1	3.38	0.06	0.34	0.02	19.99	0.21

NT Northern Territory, QLD Queensland, WA Western Australia

## Heritabilities

Heritabilities were estimated for each fruit quality trait at each site by year combination (Hardner et al. 2012) (Table 5). The highest heritabilities were associated with mesocarp recovery and average fruit weight,

indicating the relative ease at transferring these traits from parents to progeny in the Australian breeding populations. The lowest heritabilities in this study were associated with traits such as skin thickness, mesocarp thickness and deformities indicating the relative difficulty in breeding for such traits.

## Genetic Correlations

The genetic stability of traits across sites and years, may be investigated in the heat maps of genetic correlations (Fig. 1), indicating how traits are influenced by genetic and environment or seasonal conditions. Genetic stability data can also be an indicator of how transferable data is from one site to another and helps when determining varietal performance in growing areas not tested. The genetic correlations between sites and year combinations were high for most traits (with the exception of skin colour and skin thickness) especially for NT and QLD, as seen by the red blocks in each heat map (Fig. 1).

Seed width, average fruit weight and mesocarp recovery showed high stability, while skin colour showed low stability, as shown by more constant or variable colours in the heat map. The traits percent blush and mesocarp texture were fairly stable within a site and relatively stable between the Northern Territory and Queensland, however, between Western Australia and the other sites there was lower genetic correlation. Mesocarp colour was generally stable between sites and years except for year one in the Northern Territory and year two in Western Australia.

## Best linear unbiased predictions

Best linear unbiased predictions (BLUPs) of breeding values were predicted for all progeny and parent cultivars for each trait, averaged over Sites and Years. The breeding values of the parent cultivars are

presented in Fig. 2. BLUPs are centred around zero, so high positive BLUPs for a trait show cultivars that are more likely to produce progeny with desirable values for that trait while high negative values will show cultivars that are likely to have progeny with the least desirable values for that trait. For example, Cultivar 11 (Irwin) has the highest breeding values for percentage blush and blush intensity while cultivar 19 (Nam Doc Mai) has the lowest breeding values for these traits. cultivar 22 (R2E2) has the highest breeding value for mesocarp recovery while cultivar 28 (Willard) has the lowest. cultivar 24 (Suvarnareka) has the highest breeding value for mesocarp colour while cultivar 20 (Padiri) has the lowest breeding value for skin background colour.

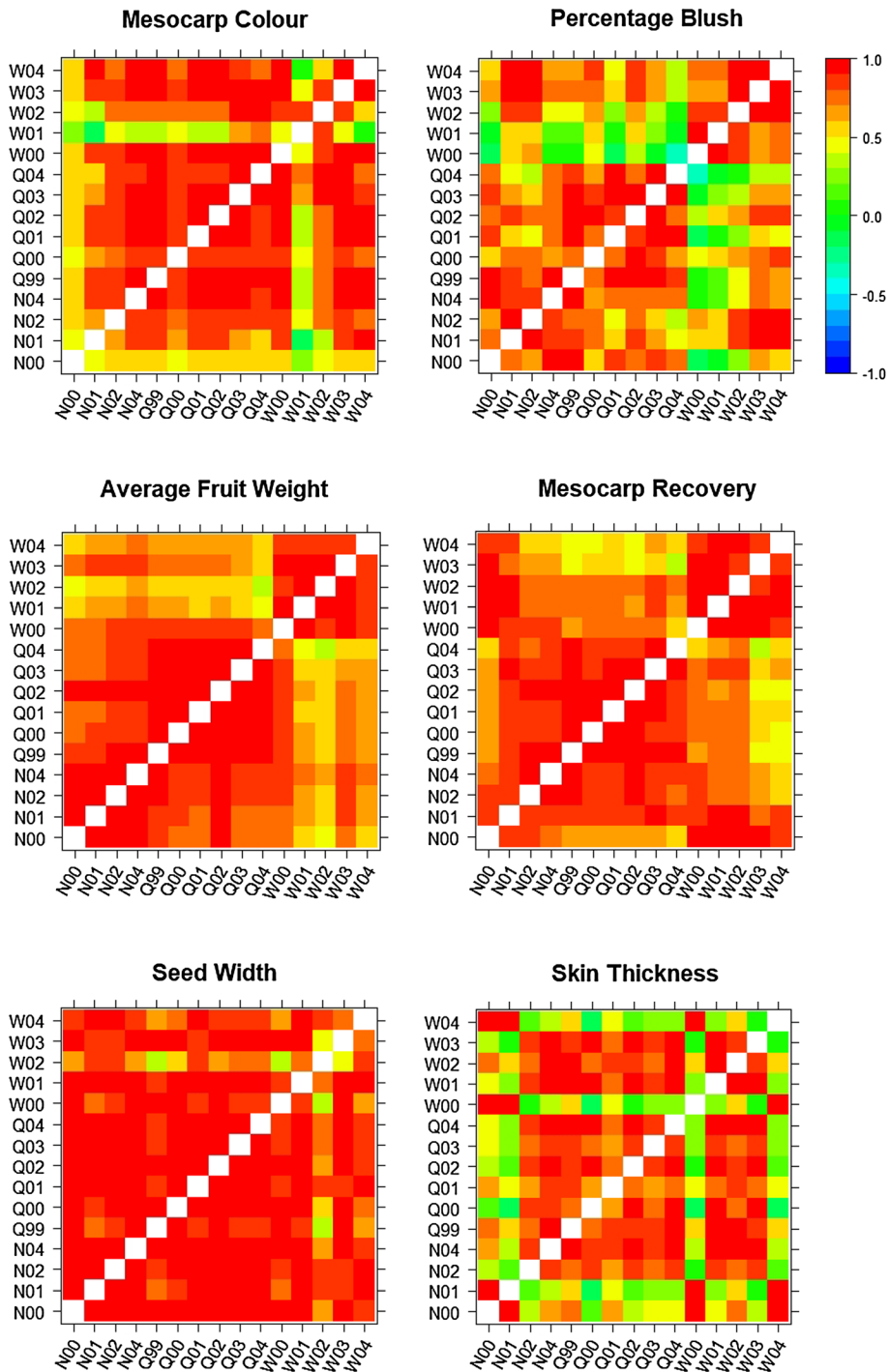
The BLUPs for each trait (averaged over sites and years) were plotted against each other in pairs (Fig. 3) to identify the predictions for parents and progeny. This plot shows the improvements in traits with high numbers of progeny showing more desirable trait values than their parents. These are identified by progeny above and to the right of parents in the top right quadrant of the scatter plots in Fig. 3. For example, there are a number of progeny with increased percent blush together with increased blush colour than any parents in the study.

To better understand the relationships among traits, the genotype BLUPs for each trait were analysed by principal component analysis (PCA). The PCA explains the variation among the traits in a smaller number of dimensions. The first two principal components were used to construct a biplot (Gabriel 1971)

**Table 5** Narrow sense heritability range and average across sites and years for 13 fruit quality traits

Trait	Heritability range	Average heritability
Average fruit weight	0.48–0.90	0.80
Mesocarp recovery	0.69–0.98	0.79
Blush intensity (no WA)	0.54–0.68	0.60
Beak shape (no WA)	0.46–0.75	0.59
Percent blush	0.32–0.93	0.57
Seed width	0.34–0.83	0.56
Blush colour	0.26–0.71	0.52
Stem-end shape (no WA)	0.26–0.82	0.49
Mesocarp colour	0.26–0.63	0.47
Skin background colour	0.27–0.65	0.47
Skin thickness	0.04–0.91	0.36
Mesocarp texture	0.13–0.99	0.35
Deformities (no WA)	0.09–0.42	0.27





**Fig. 1** Heat map representation of the genetic correlations among the site by year combinations for each fruit quality trait. The colours show the range of correlations from high positive (1.0) in red to high negative (– 1.0) in dark blue

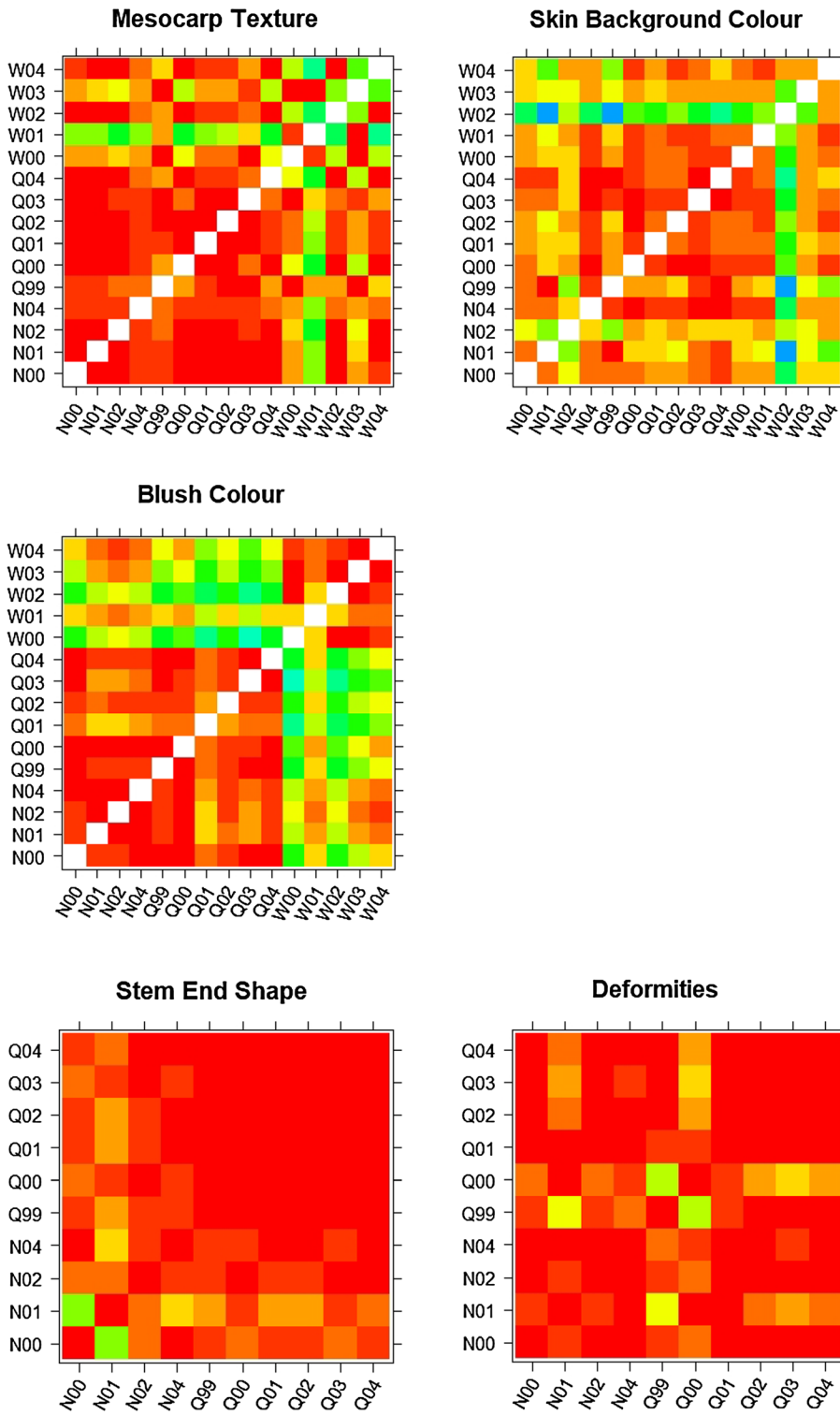
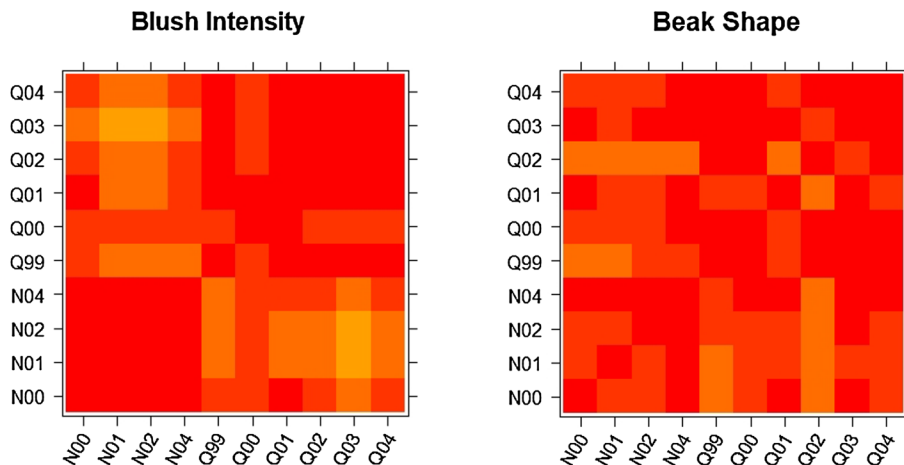


Fig. 1 continued



**Fig. 1** continued

(Fig. 4). The biplot gives an indication of how the traits are correlated, with vectors (arrows) pointing in the same direction being highly positively correlated and those in the opposite direction being highly negatively correlated. Those vectors perpendicular to each other are uncorrelated. The angle between the vectors reflects the degree of correlation between the traits with the smaller the angle the higher the correlation. Vectors extending furthest from the centre of the biplot identify variables that explain most of the variation in the data.

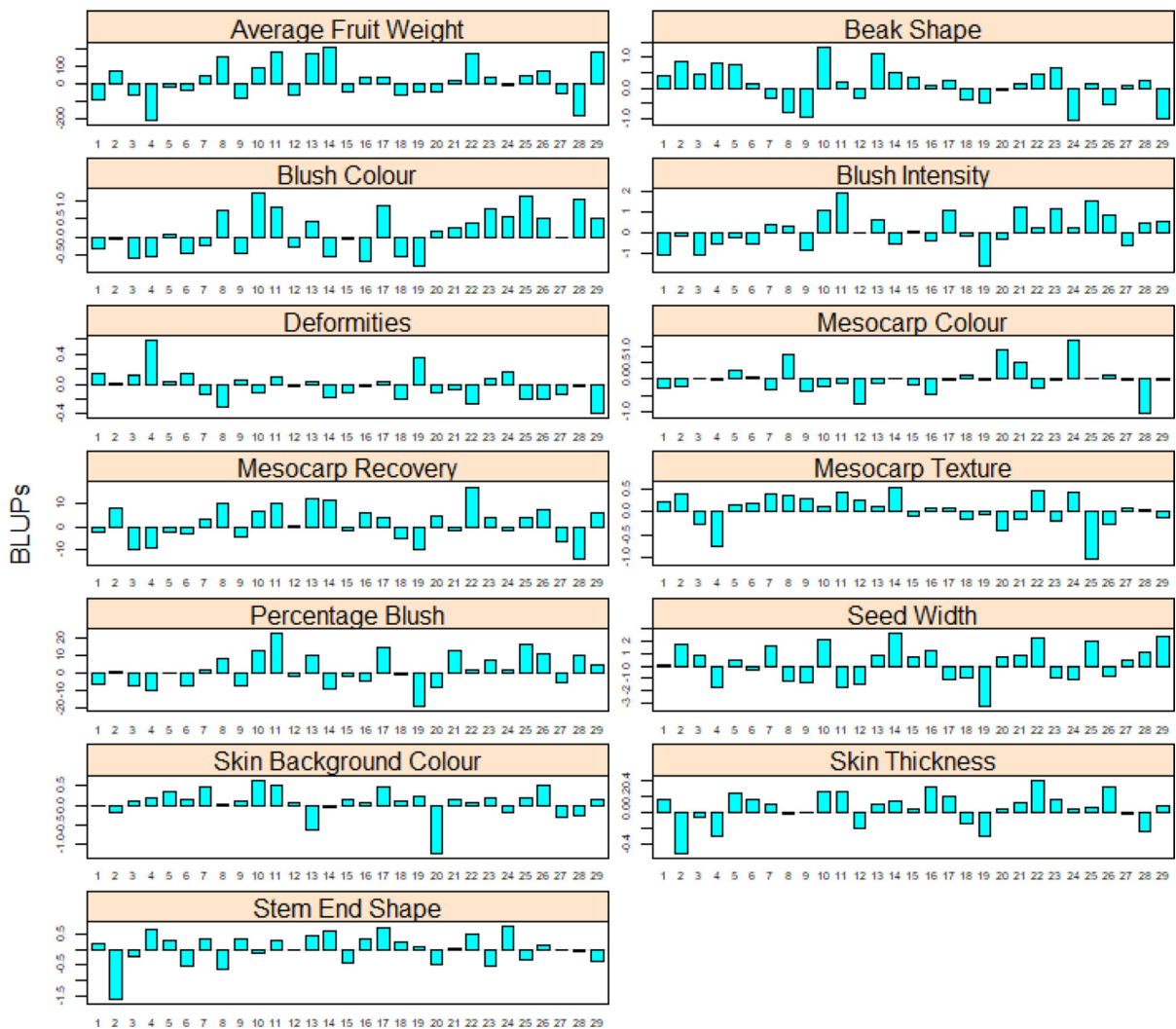
## Discussion

*Average fruit weight* is an important trait as it dictates the number of fruit that fit into each 7 kg box for marketing and very large or very small fruit are often discounted on the wholesale market. Mean average fruit weight varied between sites and between seasons (Table 4) however the genetic correlations for this trait were high (Fig. 1) The genetic correlation of average fruit weight was stable across sites and seasons, although less so in Western Australia, indicating high genetic and less environmental control over this trait and observations are relatively transferable from one site to another which is helpful in determining varietal performance in growing areas not tested. Average fruit weight had the highest average heritability (0.80) with a spread of 0.48 to 0.90 indicating the relative ease at transferring this trait from parents to progeny in the Australian breeding population. Previous estimates of

the heritability of average fruit weight by Hardner et al. (2012), were between 0.69 and 0.94. Average fruit weight for mango seems to be at the higher end of published heritabilities in fruit species for example; Japanese pear (*Pyrus pyrifolia* Nakai) at 0.73 (Abe et al. 1995), peach at 0.20 (Hansche 1986), Olive (*Olea europaea* L.) between 0.17 and 0.28 (Zeinanloo et al. 2009). There was a strong positive correlation between the average fruit weight and mesocarp recovery (Fig. 4), indicating larger fruit generally have a higher percentage of edible mesocarp and these traits can be co-selected.

*Mesocarp colour* was generally stable across sites and years with heritabilities ranging from 0.26 to 0.63, indicating it is under moderate genetic control. In Western Australia, in 2001–2002, and Northern Territory 2000–2001 the mesocarp colour correlated slightly less well with other sites and seasons. The reason for this is unclear but may be due to the stage of fruit ripeness at time of assessment. The BLUP analysis has indicated the best parental cultivars used in the study populations for improving (darkening) mesocarp colour are Padiri, Palmer, Suvarnareka and Hybrid 10. The principal component biplot shows mesocarp colour accounts for very little of the variation in the data but it seems to be positively correlated with the other mesocarp traits (mesocarp recovery, mesocarp texture) and average fruit weight. (Fig. 4).

*Mesocarp texture*, a measure of two traits (mesocarp firmness and mesocarp fibre) is a fruit trait that changes over time as fruit ripen, and as such is highly

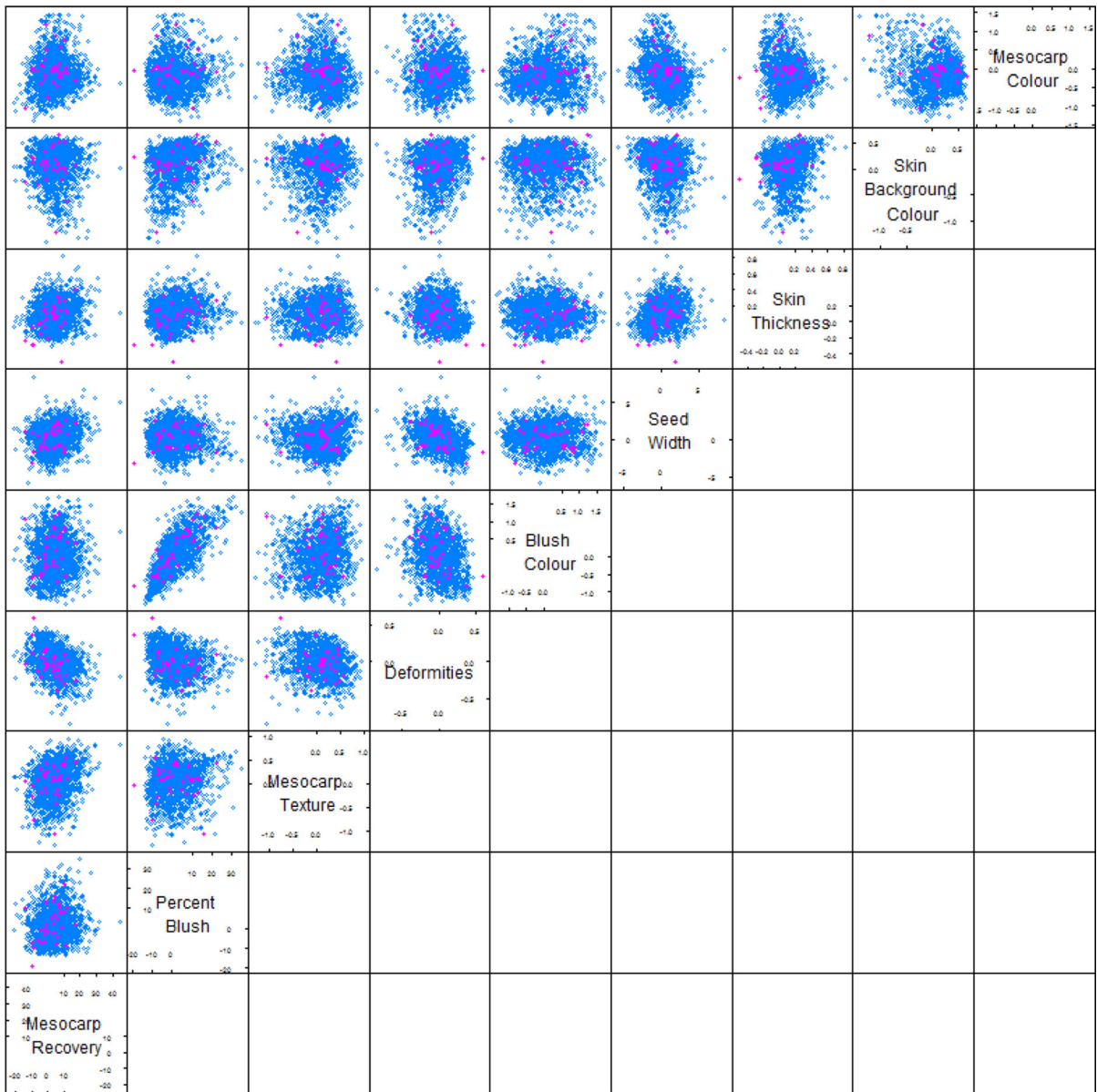


**Fig. 2** Best linear unbiased predictions (BLUPs) for each trait in each of the parent cultivars. Parental cultivars are represented on the horizontal axis by their numerical codes presented in Table 1. Predicted BLUP's are displayed on the vertical axis of each plot

influenced by the stage of fruit ripeness, which may be a reason that mesocarp texture had one of the lowest average heritabilities in this study (0.35), indicating a relative difficulty in transferring this trait between generations in a breeding program. Mesocarp texture was stable across seasons in the Northern Territory and Queensland but less so in Western Australia (Fig. 1). The principal component analysis indicated positive correlation between mesocarp texture, mesocarp recovery and average fruit weight. Separate measurement and analyses of mesocarp texture components such as firmness, fibre abundance, and fibre strength may identify which components have higher

heritability and are more useful for breeders interested in improving mesocarp texture.

*Mesocarp recovery* had one of the highest average heritability (0.79) and a high range of heritabilities (0.69–0.98), indicating the relative ease at transferring this trait from parents to progeny in the Australian breeding population. The high heritability also indicates high genetic influence and low environmental influence on the trait. We could expect the mesocarp recovery of a cultivar to be similar when grown at different sites and years making it a stable trait. From the data presented in Fig. 2, there are several parental cultivars with relatively high BLUPs indicating a



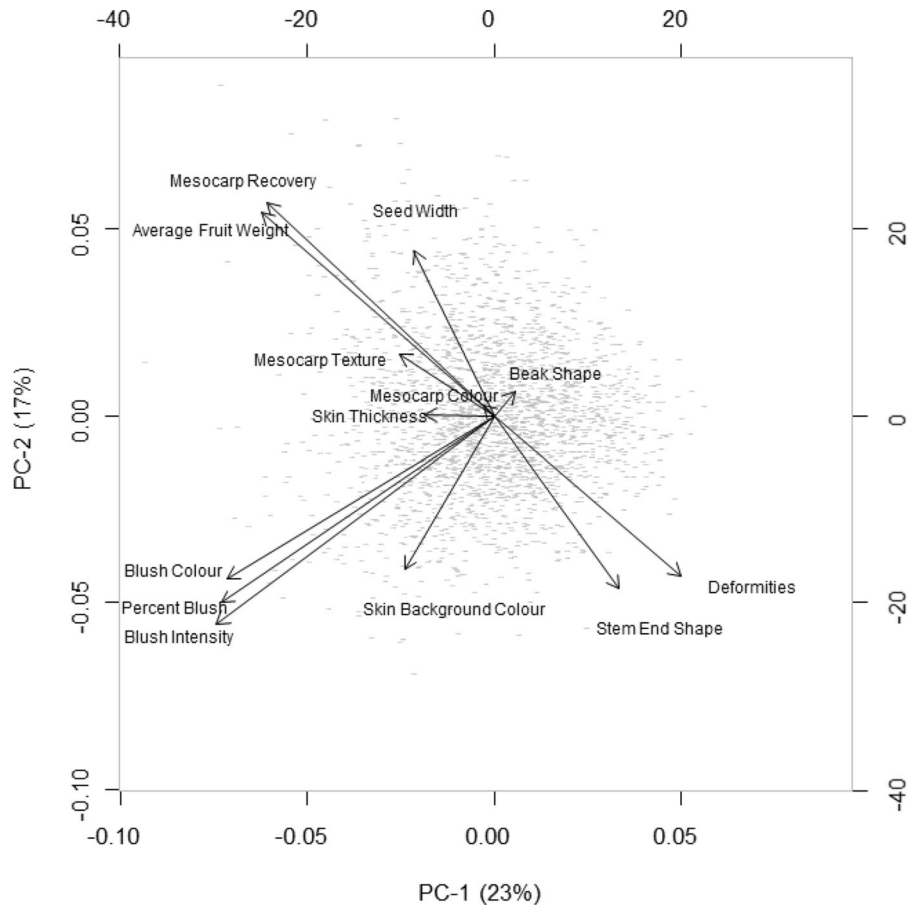
**Fig. 3** Plot of BLUPs for each pair of traits averaged over sites and years showing parents (pink) and progeny (blue)

range of parents have good ability to improve mesocarp recovery in their progeny. There was strong positive correlation between mesocarp recovery and average fruit weight (Fig. 4), indicating larger fruit generally have a higher percentage of edible mesocarp and these traits can be co-selected.

*Seed width* was another trait with high stability across sites and seasons indicating that it is mainly governed by genetics with little environmental influence on its expression. Low seed width is desirable as

it allows for more of the edible fruit mesocarp. Cultivars such as Nam Doc Mai, and perhaps Irwin, have very low BLUPs (Fig. 2) indicating that they are good parents to use to reduce seed width in hybrid progeny. The principal component analysis indicates that seed width is moderately positively correlated with mesocarp recovery and average fruit weight.

*Skin thickness* had one of the lowest average heritabilities in this study (0.36), indicating a relative difficulty in transferring this trait between generations



**Fig. 4** Biplot based on principal component analysis of genotype BLUPs from analyses of 13 fruit traits. Traits with arrows pointing in the same direction are positively correlated

and more easily co-selected whilst arrows pointing in opposite directions are negatively correlated

in a breeding program. Skin thickness was generally stable among sites and most years, however, in some site by years the relationship was poor (Fig. 1). The reason that some sites had poor correlations in some seasons is unclear but may be because of large environmental anomalies. Skin thickness was not particularly highly correlated to any other fruit quality trait (Fig. 4).

*Skin background colour* had a reasonably large spread of heritabilities (0.27–0.68, av. 0.47) (Table 5) and BLUPs (Fig. 2) indicating the importance of parental selection when breeding for this trait. The Floridian parent cultivars Van Dyke, Irwin, Haden and Lippens had the strongest BLUPs and likelihood of transferring this trait to progeny. Across sites and years, skin background colour was less stable than mesocarp colour and texture, indicating a higher

environmental influence on this trait. The de-greening of mango skin during ripening can depend on the nitrogen status of the fruit and ripening temperatures (Hofman 1997). The biplot (Fig. 4) shows fruit skin background colour is moderately positively correlated with blush colour, percent blush and blush intensity.

*Blush colour* in many Asian cultivars are unblushed whereas those originating from Florida are often highly-blushed. In Australian and other markets, blushed fruit receive a premium price due to their eye appeal. Heritability of blush colour varied from 0.26 to 0.71 with an average of 0.52 indicating a reasonable ease in transferring this trait to progeny when the best parents are used. The genetic correlations in the studied populations show blush colour was stable within sites and across sites in Queensland and the Northern Territory but not in Western

Australia. The reasons for the difference in blush colour in Western Australian are unclear but may have been due to the relative difference in fruit ripeness or differences in tree shading and light transmission due to differences in pruning between sites. Blush development requires fruit skin exposure to direct sunlight (Berardini et al. 2005a, b). As might be expected, the skin blush colour traits (percent blush, blush colour and blush intensity) were strongly positively correlated indicating all three can be bred for or selected simultaneously.

*Percentage of blush* covering the skin has similar variability to blush colour within sites and was relatively stable between Queensland and the Northern Territory but not in Western Australia. The reasons for the difference in the percentage of skin covered with blush in Western Australian fruit are unclear but may also have been due to the relative difference in tree shading and light transmission due to differences in pruning between sites. The percentage of skin covered by blush is negatively influenced by shading within the tree and as such can be managed through pruning and training of canopies.

*Blush intensity* had a reasonably high average heritability (0.6) and was stable across years and sites (only data from Queensland and Northern Territory used here). The parental cultivar Irwin had high BLUPs for Blush Intensity. The principal component biplot showed high positive correlation between blush intensity, blush colour and percent blush.

*Stem-end* shape had relatively large range and medium average heritability (0.26–0.82, av 0.47) in this study. Stem-end shape was strongly negatively correlated with average fruit weight (Fig. 4) which may be contributing to the large range of heritabilities. Stem-end shape was strongly positively correlated with deformities (Fig. 4) indicating the stem-end of the fruit may be influencing the level of fruit deformities.

*Fruit deformities* had one of the lowest heritabilities (0.09–0.42, av.0.27) in this study (Table 5), indicating the low genetic component and relative difficulty in breeding for such trait. Fruit deformities are often caused by environmental conditions such as excessive temperatures or nutritional deficiencies during fruit development. Fruit deformities are highly negatively correlated with mesocarp recovery and average fruit weight, indicating that heavier fruit are more likely to have less deformities. There was also a strong positive

correlation between deformities and the stem end shape of the fruit, as growth deformities often occur at the stem end of the fruit.

#### Comments on method and limitation of study design

The statistical analysis approach presented here has successfully modelled mango genetic effects for multiple traits, several years and multiple environments, allowing insight into the heritability and stability of traits and the relationship among traits across environments. The linear mixed model, incorporating pedigree information and modelling genotype by environment effects using factor analytic models, provides a comprehensive multivariate modelling approach, however there are limitations in this study. Firstly, the sparse data on trees across sites and years has made modelling spatial and temporal correlation problematic and only simple residual models have been able to be fitted. In other studies in perennial crops, spatial and temporal correlation has been found to be significant (Stringer and Cullis 2002; Dutkowski et al. 2002; Smith et al. 2007; De Faveri et al. 2015) and so the simple residual models fitted in this paper may not be optimal. However as different trees were measured at different times the effect on predictions may not be large. Also, the limited numbers of progeny per parent create a very sparse unbalanced data set which may have been improved with more data on more crosses.

A similar mixed model approach was implemented in the univariate analysis of mango fruit weight (Hardner et al. 2012). In that paper only factor analytic models with a single factor (FA1) were fitted and hence were not found to be the best model. In our case we have fitted higher order factor analytic models (with two to four factors) and in all cases the higher order factor analytic models were a significant improvement on a single factor (FA1) model. The factor analytic model allows a good approximation to the fully unstructured covariance model where all variances and pairs of covariances are estimated, but it is important to fit sufficient factors for accurate separation of genetic and non-genetic effects. Failure to fit sufficient factors will result in biased estimates of genetic effects due to interplay between genetic and residual components in the model (De Faveri 2017).

The stability analysis of traits across sites and years as shown in the heatmaps (Fig. 1) allows insight into which traits may be combined across sites and which may require more specific environment by year testing in future studies. In general, the analyses showed the Northern Territory and Queensland sites to be very similar with high genetic correlations between these sites for most traits (for example average fruit weight, mesocarp recovery, seed width, blush intensity), hence selection based on one of these sites is likely to correspond favourably with the other. The Western Australian site showed some differences to these two sites with lower correlation for traits such as mesocarp texture and blush colour. Traits such as mesocarp colour, skin background colour and skin thickness showed differences both among sites and among years within sites, and may need more intense sampling. Most other traits showed very high genetic correlation between years within a site and hence may not need to be sampled every year.

The issue of how best to obtain genetic parameters from tree breeding programs where data needs to be sampled across sites, years, traits and family groups when not all trees can be measured for all traits at all times due to time and labour constraints is an interesting topic of future research. Optimal sparse sampling designs could be developed to optimize the accuracy of genetic parameter prediction. Having more genetic information, for example, genomic marker data on the trees would also improve the power and estimation of genetic effects in sparse designs.

The approach implemented here analyses each trait individually using the linear mixed model and then the BLUPs from each analysis have been subjected to a principal component analysis in order to investigate relationships among traits. A full multi-trait, multi-year, multi-site analysis would have been preferable to estimate genetic correlations among traits, however the computational burden for such an analysis was prohibitive.

## Conclusion

The analyses presented here on fruit quality traits have improved our understanding of their heritability and the relative ease of difficulty of transferring these traits from parents to progeny in a controlled hybridisation

program. The findings on the stability of these fruit quality traits across years and environments will help in designing future regional performance trials and in predicting performance in other non-tested environments. The principal component analyses and visual representation in the biplot presented in Fig. 4 has highlighted where certain fruit quality traits are closely correlated, indicating that selection of ideal parents for one of these traits is likely to deliver progeny that also have higher representation of the other highly correlated trait. Interpretation of heritability estimates and results from other analyses presented in this report must be done in the context of the populations used in the analysis. Relationships and heritabilities may change in populations of other breeding programs with different parents and genetic profiles.

**Acknowledgements** We would like to acknowledge the efforts of Peter Johnson, David Hamilton, Chris Wicks, Ping Lu, Richard Brettell, Sam Blakie, Warren Muller and Craig Hardner who contributed to collecting and collation of the phenotypic data used in this analysis. We would like to acknowledge the Queensland Department of Agriculture and Fisheries, the Northern Territory Department of Primary Industry & Resources, the Western Australian Department of Primary Industries and Regional Development and CSIRO for their financial support of the National Mango Breeding project.

**Author contributions** Both authors contributed to the study conception and design. Material preparation, and data collection was performed by Ian Bally and data analyses were performed by Joanne De Faveri. The first and subsequent drafts of the manuscript were written by Ian Bally and Joanne De Faveri. All authors read and approved the final manuscript.

**Funding** The National Mango Breeding project was supported by the Queensland Department of Agriculture and Fisheries, the Northern Territory Department of Primary Industry & Resources, the Western Australian Department of Primary Industries and Regional Development and CSIRO.

**Data Availability** The datasets generated analysed during the current study are not publicly available due to commercial in confidence and the protection of intellectual property associated with the Australian National Mango Breeding Program but are available from the corresponding author on reasonable request.

**Code availability** Statistical methods and software used in this manuscript are outlined in the materials and methods section.



## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Abe K, Sato Y, Saito T, Kurihara A, Kotobuki K (1995) Narrow-sense heritability of fruit characters in Japanese pear (*Pyrus pyrifolia* Nakai). *Jpn J Breed* 45(1):1–5. <https://doi.org/10.1270/jsbbs1951.45.1>
- AMIA (2014) Mango Strategic Investment Plan 2014/15. AMIA, p 25
- Bally IE, Lu P, Johnson P (2009) Mango breeding. In: Jain SM, Priyadarshan PM (eds) *Breeding plantation tree crops: tropical species*. Springer, New York
- Bally ISE (2008) Mango breeding activities in Australia. Paper presented at the 4th international symposium on tropical and subtropical fruits, Bogor, West Java, Indonesia, 3–7 November
- Bally ISE, Akem CN, Dillon NL, Grice K, Lakshesar D, Stockdale K (2013) Screening and breeding for genetic resistance to anthracnose in mango. *Acta Hort* 992:239–244
- Bally ISE, De Faveri J, Maddox C (2017) Breeding for low vigour in mango. *Acta Hort* 1138:63–68. <https://doi.org/10.17660/ActaHortic.2017.1183.9>
- Bally ISE, Dillon NL (2018) Mango (*Mangifera indica* L.) Breeding. In: Al-Khayri JM, Jain SM, Johnson DV (eds) *Advances in Plant Breeding Strategies: Fruits*, vol 3 *Advances in Plant Breeding Strategies*. Springer International Publishing, Cham
- Bally ISE, Lu P, Johnson P, Muller WJ, González A (2009) Past, current and future approaches to mango genetic improvement in Australia. *Acta Hort* 820(1):153–163
- Berardini N, Fezer N, Conrad J, Beifuss U, Carle R, Schieber A (2005) Screening of mango (*Mangifera indica* L.) cultivars for their contents of flavonol O- and xanthone C-glycosides, anthocyanins and pectin. *J Agric Food Chem* 53(5):1563–1570
- Berardini N, Schieber A, Kläiber I, Beifuss U, Carle R, Conrad J (2005) 7-O-methylcyanidin 3-O-b-D-galactopyranoside, a novel anthocyanin from mango (*Mangifera indica* L. cv. ‘Tommy Atkins’) peels. *Z Naturforsch* 60:801–804
- Bernardo R (2010) *Breeding for Quantitative Traits in Plants*. Stemma Press, Woodbury, Minn
- Butler DG, Cullis BR, Gilmour AR, Gogel BJ (2009) ASReml-R reference manual, release 3. Department of Primary Industries and Fisheries, Brisbane, Australia, Toowoomba, Queensland
- Cullis BR, Smith AB, Beeck CP, Cowling WA (2010) Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. This article is one of a selection of papers from the conference exploiting genome-wide association in oilseed brassicas: a model for genetic improvement of major OECD crops for sustainable farming. *Genome* 53(11):1002–1016. <https://doi.org/10.1139/G10-080>
- De Faveri J (2013) Spatial and temporal modelling for perennial crop variety selection trials., PhD thesis, University of Adelaide, Adelaide
- De Faveri J, Verbyla AP, Pitchford WS, Venkatanagappa S, Cullis BR (2015) Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. *Crop Pasture Sci* 66(9):947–962. <https://doi.org/10.1071/CP14312>
- De Faveri J, Verbyla AP, Lee SJ, Pitchford WS (2017) Maternal body composition in seedstock herds: 5. *Animal Production Science, Multivariate analysis using factor analytic models and cluster analysis*. <https://doi.org/10.1071/AN15465>
- Dutkowski G, Costa e Silva J, Gilmour A, Lopez G (2002) Spatial analysis methods for forest genetic trials. *Canadian Journal of Forest Research* 32:2201–2214
- Falconer DS, Mackay TFC (1996) *Introduction to Quantitative Genetics*. Longman Scientific and Technical, London
- Gabriel KR (1971) The biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58:453–467
- Galan Saco V (2017) Trends in world mango production and marketing. *Acta Hort* 1183:351–364
- Hansche PE (1986) Heritability of fruit quality traits in peach and nectarine breeding stocks dwarfed by the dw gene. *Hortsci* 21(5):1193–1195
- Hardner C, Bally I, Wright C (2012) Prediction of breeding values for average fruit weight in mango using a multivariate individual mixed model. *EUP* 168(2):463–477. <https://doi.org/10.1007/s10681-012-0639-7>
- Hofman PJ (1997) What causes green, ripe mangoes? *Mango Care Newsletter* 20:13–15
- Horticulture Innovation Australia (2017) *Mango Strategic Investment Plan 2017–2021*. Hort Innovation, Sydney, Australia
- Horticulture Innovation Australia (2018) *Australian Horticulture Statistics Handbook, Fruit, 2016/2017*. Horticulture Innovation Australia, Sydney, Australia
- Iyer CPA, Schnell RJ (2009) Breeding and Genetics. In: Litz RE (ed) *The Mango Botany, Production and Uses* 2nd Edition, 2nd edn. CABI International, Wallingford, Oxen, pp 67–96
- Kelly AM, Smith AB, Eccleston JA, Cullis BR (2007) The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Sci* 47:1063–1070. <https://doi.org/10.2135/cropsci2006.08.0540>
- Kulkarni VJ, Bally ISE, Brettell RIS, Johnson PR, Hamilton D (2002) The Australian national mango breeding program - in search of improved cultivars for the new millennium. *Acta Hort* 575(1):287–293
- Medlicott AP, Bhogal M, Reynolds SB (1986) Changes in peel pigmentation during ripening of mango fruit (*Mangifera indica* var. Tommy Atkins). *Ann Appl Biol* 109:651–656
- Meyer K (2007) Multivariate analyses of carcass traits for Angus cattle fitting reduced rank and factor-analytic models. *J Anim Breed Genet* 124:50–64. <https://doi.org/10.1111/j.1439-0388.2007.00637.x>
- Pott I, Marx M, Neidhart S, Muhlbauer W, Carle R (2003) Quantitative determination of beta-carotene stereoisomers in fresh, dried, and solar-dried mangoes (*Mangifera indica* L.). *J Agric Food Chem* 51(16):4527–4531

- Proctor JTA, Creasy LL (1969) The anthocyanin of the mango fruit. *Phytochemistry* 8(10):2018
- R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/>. 2017
- Smith AB, Cullis BR, Thompson R (2001) Analysing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics*. <https://doi.org/10.1111/j.0006-341X.2001.01138.x>
- Smith AB, Stringer JK, Wei X, Cullis BR (2007) Varietal selection for perennial crops where data relate to multiple harvests from a series of field trials. *Euphytica* 157:253–266
- Stringer JK, Cullis BR (2002) Application of spatial analysis techniques to adjust for fertility trends and identify interplot competition in early stage sugarcane selection trials. *Aust J Agric Res* 53:911–918
- The Royal Horticultural Society (2001) RHS colour chart, 4th edn. Royal Horticultural Society, London
- Zeinanloo A, Shahsavari A, Mohammadi A, Naghavi MR (2009) Variance component and heritability of some fruit characters in olive (*Olea europaea* L). *Sc Hort* 123(1):68–72. <https://doi.org/10.1016/j.scienta.2009.07.024>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.