

Canola yield and its association with phenological, architectural and physiological traits across the rainfall zones of southwestern Australia

Heping Zhang, Jens Berger, Chris Herrmann, Adam Brown, Sam Flottmann
CSIRO Agriculture and Food, Private Bag 5, Wembley, WA6014

Abstract

Canola yield is a complex trait determined by environmental and genetic factors, and their interaction. We investigated yield performance of 21 canola across multiple environments and its association with flowering time, architectural and physiological traits. Pattern analyses showed that environment discriminated varieties differently, from which two mega-environments and four variety groups were identified. Principal component analysis (PCA) revealed specific adaptation: i.e. yield was associated with early flowering in the low/medium rainfall zone (LMRZ) and with late flowering in the high rainfall zone (HRZ). A few varieties outperformed all other varieties across environments and were broadly-adapted. PCA revealed that yield was positively associated with biomass and HI but negatively with days to flowering in the LMRZ, and positively with biomass, longer days to flowering and more seeds m^{-2} but negatively with HI in the HRZ. The combination of high harvest index and biomass of broadly-adapted varieties enabled them to produce high yield across the environments.

Introduction

Canola yield is a complex trait determined by both environmental and genetic factors, and their interaction (Zhang *et al.* 2013; Zhang *et al.* 2017). Biomass, siliquae per plant and thousand-seed weight are reported to be the major contributors to high yield. However, the relationships between yield and phenological, architectural and physiological traits in canola have not been studied in depth and how the phenological, architectural and physiological traits interact with the environment are poorly understood. The indeterminate canola crop produces extensive above-ground vegetative biomass and hence has a relatively low harvest index compared with the determinate crops such as wheat. Understanding the relationships between yields with these traits across environments will help breeders to focus on traits leading to yield improvement and growers to manage the crop better. Identification of traits contributing to high yield and yield stability can help prioritise breeding strategies in selecting both specific and broad adapted genotypes to achieve greater yield gains.

The major objective of this study was to i) understand the adaptation and yield formation of canola in the low, medium and high rainfall zones of WA; ii) use principal component analysis (PCA) to establish the relationships between yield and the phenological, architectural and physiological traits of a set of canola varieties/lines; iii) identify the phenological, agronomic and physiological traits that contribute to high yield and yield stability in different environments.

Materials and methods

We conducted a total of 10 experiments at the low (Merredin: MR), medium (Cunderdin: CD) and high (Kojonup: KJ) rainfall zones of WA from 2013 to 2017. These experiments involved 20-37 open-pollinated (OP) and hybrid Triazine tolerant (TT), Clearfield® (CL), conventional (CV) and Roundup Ready® (RR) canola with contrasting phenology (early, mid and late flowering), early vigour (hybrid vs OP), and yield potential (low and high). Treatments varied among site and years but consistently followed the genotype by environment by management ($G \times E \times M$) approach. The treatments consisted of low (0 kg/ha) and high (80-150 kg/ha) N rate, rainfed and irrigated, early and late sowings. The experiments were replicated three times at each site with a split-plot design with the herbicide tolerance groups allocated to the main plot and variety in the subplot. The plot size was 20m by 1.54m at all sites. We selected the high yielding treatment at the ten environments for analysing yield and its association with phenological, architectural and physiological traits. To save space, we adopt an abbreviation of the site followed by year, and treatment to represent the individual environment. The ten high yielding environments are KJ2013H, KJ2014H, KJ2015R, KJ2016L and KJ2017E in the HRZ and CD2013L, CD2014H, MR2014L, MR2014H, MR2015E and MR2016E in the LMRZ. We recorded the day of flowering, plant height, lowest branch height, pod canopy depth, branch number, biomass and harvest index (HI), yield and yield components (pods/plant, pods/ m^2 , thousand seed weight (TSW), seeds/pod). Yield was calculated based on the seed weight from the whole plot. Oil content, protein and moisture content were measured using a NIR grain analyser. Yield was standardized and reported at 8% of moisture and 42% of oil content for all varieties. At maturity, three plants from each plot were harvested for counting the number of pods per plant and seeds per pod. Biomass and HI were measured using a quadrat of 0.55 m^2 of quadrat samples at maturity.

We used the ASReml-R package to fit the mixed models for the multiple environment trials, estimated the variance parameters, and predicted the fixed and random effects for yield, biomass, seeds m^{-2} , TSW and HI

traits. The environment was treated as fixed effect; genotype and the G×E interaction as random effect. The best linear unbiased prediction (BLUP) from the model was used to predict trait values. These predicted trait values were used for classification analysis of varieties and environments, yield-traits association analyses and principal component analysis (PCA). For classification, a hierarchical clustering method was applied to the matrix of yield and the squared Euclidean distance was used as the dissimilarity measure. Dendrograms for yield were constructed to investigate similarities in the pattern of performance among varieties in response to the environmental changes and environments in discriminating varieties. To investigate the G×E interaction and the relationships among the traits, we conducted two types of principal component analysis. The first one was used to identify the G×E interaction and reveal the interrelations among genotypes and environments. The second one is to quantify the relationships among yield, phenological, architectural and physiological traits. We used the FactoMineR package under the R environment to implement PCA based on the predicted mean values of traits from the FA model.

Results

PCA analysis reveals that the environments discriminated the varieties differently and the first two principal components accounted for 92% of the G×E interactions for yield (Fig. 1). The angles between the HRZ environment vectors (KJ2013H, KJ2014H, KJ2016E, KJ2017E) and LMRZ environment vectors (MR2014L, MR2015E, CD2013L and CD2014H) are greater than 90°, suggesting there were two different mega-environments. This was confirmed by hierarchy cluster analysis (Fig. 2a). The first mega-environment consisted of KJ2013H, KJ2014H, KJ2016E, KJ2017E and MR2016E, representing HRZ environment and early sowing in the LRZ environment. The second mega-environment included MR2014L, MR2015E, KJ2015R, CD2013L and CD2014H, representing the LMRZ environments along the below-average rainfall environment in HRZ.

Variety performance was different between the two mega-environments (Fig. 1). Cluster analysis of yield showed that the varieties could be divided into four groups (Fig. 2b). Variety Group 1 consisted of five hybrid varieties (NS Diamond, H404RR, P43Y23RR, H600RR and P45Y25RR), showing good performance across all ten environments (Fig. 1). The first three varieties were early flowering hybrids, closely associated with the LMRZ environments (MR2015E, MR2014L, CD2014H, CD2013L), and produced the highest yield in the LMRZ environments. The last two were late flowering, more associated with the HRZ environments, producing the highest yield in HRZ and comparable yield to Group 3 and 4 in LMRZ environments. The high yield of Group 1 was from high biomass, above-average HI, more seeds m⁻² and average TSW (data not shown). These varieties appeared to have broad adaptability because they were either the highest yielding varieties or produced yield close to the highest yielding varieties. Group 2 consisted of eight OP and hybrid varieties (H635CC, H50, H559TT, ATR Stingray, ATR Wahoo, H750, ATR Bonito, and H577CL, mostly mid-flowering). This group was closely associated with the HRZ environment vectors (KJ2016L, KJ2017E, KJ2014H, and KJ2015R) (Fig. 1a), produced more seeds m⁻² and consequently higher yield from high biomass but lower HI (Figs. 3a, b). This group of varieties located at the opposite direction of LMRZ environment vectors (MR2015E, MR2014L, CD2014H, CD2013L) (Fig. 1a), being negatively associated with the LMRZ environments and producing low yield from lower biomass (Figs. 3c, d). Variety Group 3 consisted of four early flowering OP and hybrid varieties (P43C80, CB Tango, GT Viper, and P44Y87). It produced similar yield to variety Group 1 and 4, higher yield than variety Group 2 in the LMRZ environments. However, it was located in the opposite direction of the HRZ environment vectors and produced the lowest yield in the HRZ environments (Fig. 1a). The associations of variety groups to environments suggest that variety Group 2 was better adapted to the HRZ environments and Group 3 to the LMRZ environments. The fourth group included four varieties (H575CL, AV Garnet, GT Cobra and P45Y86) located close the origin of the biplots, indicating their average performance across the environments.

PCA analysis revealed that the first two components explained 52-63% of the total variance of the phenological, architectural and physiological traits in the HRZ environments and 49-70% in the LMRZ environments (Fig. 3). The third component explained a further 10-20%. In the PCA biplots, the individual trait vectors and the points represented four variety groups were used to facilitate visualization of the relationships among the traits. The correlation coefficient between any two traits is approximated by the cosine of the angle between their vectors. The vectors with angles near 0° are positively correlated whereas angles near 180° indicate negatively correlated. An angle close to 90° indicates a near-zero correlation.

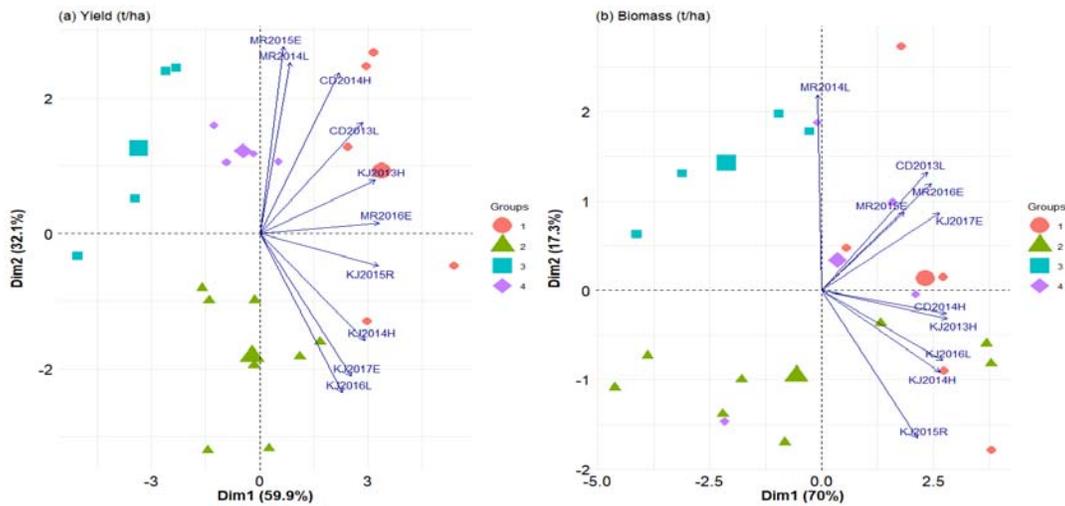


Fig. 1 Biplot of the first and second principal components for yield and biomass of 21 canola varieties in 10 environments across the rainfall zones of Western Australia. Variety groups can be referred to the text.

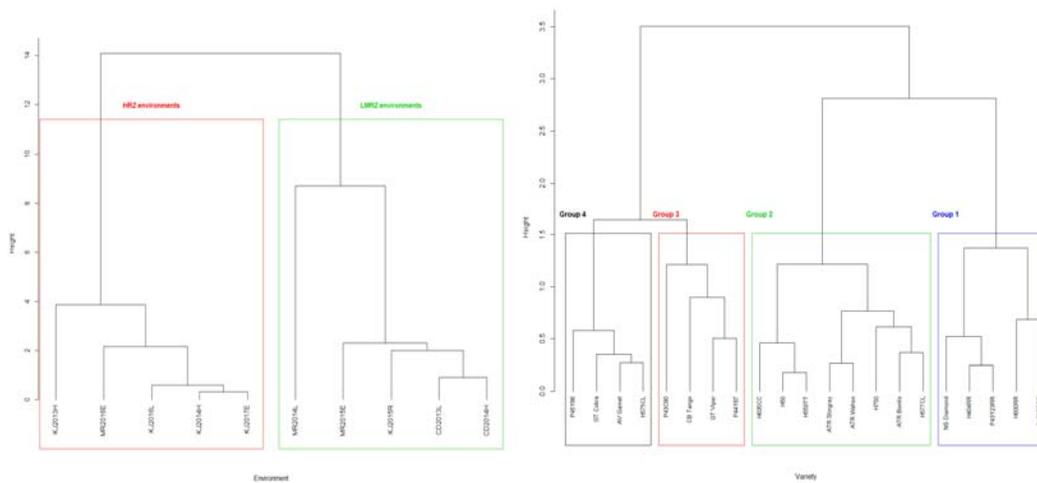


Fig. 2 Hierarchy for classification 10 environments and varieties across the rainfall zones of Western Australia and based on the yield difference of 21 canola varieties.

In the majority of the environments (KJ2013H, KJ2015R, CD2013H, CD2014H, MR2014L, MR2015E), PC1 accounted for more 50% of yield variation. However, PC2 accounted for 78% of yield variation at MR2016E, and PC2 (41%) and PC3 (35%) for a similar amount of variation at KJ2016L. The traits contributing to PC1, PC2, and PC3 are different between the LMRZ and HRZ environments. In the HRZ, PC1 represents phenology-productivity traits; PC2 represents yield components and the efficiency of converting biomass to reproductive organs (pods and seeds) and yield (HI); The contribution of architecture traits (pod canopy thickness, pods per plant and the lowest branch height) was mainly on PC3. The acute angles among the yield vector and vectors of biomass, days to flowering, seeds m^{-2} indicate that yield was positively correlated to these traits (Figs. 3a, b). The large obtuse angles among the HI, biomass, and day to flowering, plant height vectors suggest that HI decreased with increased plant height and biomass and delay in flowering. The association between yield and biomass in the LMRZ environments was similar to that in the HRZ environments as indicated by the acute angles between their vectors (Figs. 3a, b). However, the association among yield, days to flowering, and HI are in contrast to that in the HRZ environment. Yield was positively associated with HI but negatively with days to flowering.

The most significant difference in the trait associations biplots among the variety groups between the LMRZ and HRZs is the scatter of variety groups in the biplots. Specifically-adapted varieties showed different association to environments mainly by phenology. For example, Group 2 varieties located in the right quadrants in the HRZ environments, but moved to the left quadrants in the LMRZ environments. In contrast, Group 3 varieties located in the third quadrant in the HRZ environment, but moved to the first quadrant in the LMRZ environment. This reflects specific adaptation of these two variety groups to HRZ and LMRZ environments, respectively. The broadly-adapted variety Group 1 was always associated with the yield vector.

