

A genomic selection training population for phalaris: genetic composition and seasonal yield

R Culvenor¹, K Smith^{2,3}, J Forster³, N Cogan³, L Pembleton³ and J Sewell⁴

¹CSIRO Agriculture and Food, GPO Box 1700, Canberra ACT 2601, richard.culvenor@csiro.au

²Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Private Bag 105, Hamilton, VIC 3300

³AgriBio, Department for Economic Development, Jobs, Training and Resources, 5 Ring Road, Bundoora, VIC 3083

⁴PGGWrightson Seeds, Ballarat Research Station, PO Box 1425, Ballarat MC, VIC 3354

Abstract

Genomic selection (GS) is a new form of marker-assisted selection which offers the potential for more efficient selection of quantitative traits such as yield in species with long selection cycles. GS requires accurate genotyping and phenotyping of a 'training population' which is then used to select on genotype in a breeding population. We describe a training population developed to test GS in phalaris and present seasonal herbage yield data from one of three evaluation sites. The population contains 290 seed-retaining, winter-active families based Holdfast GT, Advanced AT, Landmaster and a Holdfast-derived population. Progeny were evaluated with commercial cultivars from 2014 -16 at Canberra for herbage DM yield in rows and a subset also evaluated in swards. A high level of phenotypic variation was observed with high correlations between observations in the main growing season within and between years. Correlations between row and sward yield were 0.7-0.8 in the growing season and lower at other times when stem material was present. The data from this and other sites will be combined with marker information to assess the feasibility of using GS in phalaris breeding.

Keywords

Forage breeding.

Introduction

The prominent role phalaris (*Phalaris aquatica* L.) has played in improved pastures in southern Australia has been underpinned by a sustained breeding effort within Australia since the 1960's. This has resulted in gains in seedling vigour, winter production, adaptation to lower rainfall areas, seed retention and acid soil tolerance. Future effort must address the continued need for productivity gains to maintain industry profitability, and evolving soil, climatic and management challenges. This effort must occur within a context of highly constrained resources for pasture breeding, less public sector involvement in cultivar development, and the opportunity for increased rate of genetic gain through efficient use of quantitative genetics and modern genomics tools.

Genomic selection (GS) is a relatively new form of marker-assisted selection suited to polygenic traits (controlled by many genes) that employs large numbers of markers across the whole genome facilitated by greatly reduced costs of genotyping (Meuwissen et al. 2001). Deriving a relationship between phenotype, e.g. yield, and allelic effects at all of these markers allows a Genomic Estimated Breeding Value to be calculated for a genotype. Once a robust prediction equation is developed, one or more cycles of selection based solely on genotyping seedlings can be undertaken before further recalibration, reducing the need for long and expensive phenotypic evaluation under conventional schemes. GS also allows the selection of individual plants for traits that can normally be evaluated only in swards. This gives the potential for within-family selection in addition to family selection which is more normal in outcrossing forage species.

The implementation of GS depends on the development of an accurate relationship between phenotype and genotype in a 'training' or 'reference' population. This relationship can then be applied to a breeding population of interest. While obligate outcrossing species in theory require large training populations (1000's of individuals) to be genotyped and phenotyped, there are some recent studies (e.g. Resende et al. 2014) which conclude that GS can still be significantly advantageous over conventional breeding with smaller training populations (100's of individuals). A training population for phalaris is described in this paper and preliminary results on seasonal yield presented for one of three sites at which the training population is being evaluated. Families were evaluated as drill rows due to resource and seed limitations but some families were also evaluated as swards.

Methods

Composition of phalaris training population

The training population is based on seed-retaining, winter-active germplasm, concentrating on the main phalaris zone (>550 mm LTA rainfall), since this offers the potential for highest gains in production, particularly during winter when extra feed is of high value. The germplasm pool represented by Holdfast and related cultivars such as Landmaster, Holdfast GT and Advanced AT was considered suitable for this purpose and incorporates gains from recent selection for persistence factors. Accordingly, the training population is based 50% on crosses among and within Holdfast GT, Advanced AT and Landmaster, 32% on crosses among progeny of Advanced AT and a related population, and 18% on a commercial population derived from crosses between Landmaster and Advanced AT with a Holdfast-related population. The resulting training population comprising 290 half-sib families should be applicable to future commercial breeding programs in the seed-retaining, winter-active phalaris pool. A partial sequence based on RNA (transcriptome) has been completed for phalaris and ~63,000 single nucleotide polymorphism (SNP) markers identified to genotype training population individuals (Baillie et al. 2017).

Evaluation trials

Progeny of the 290 parent plants are being evaluated in replicated rows at Canberra (ACT), Hamilton and Maryborough (Victoria). Only data from Canberra are presented here. Since DM yield assessed in rows will be correlated with yield in swards, a subsample of 33 cultivars and families was also grown as swards to quantify this correlation.

A total of 305 entries (290 families plus 15 control lines) were sown in rows spaced 40 cm apart in May 2014. Row length was 4 m and there were 3 replicates in a row-column design. Sowing rate was equivalent to 4.5 kg/ha of viable seed. Weeds were controlled by handweeding or broadleaf sprays as required. DM yield was evaluated in 2015 and 2016 by calibrated visual assessment using 2 observers 5 times per year, around late May/early June (autumn yield), late August (winter yield), mid-October (main spring growth), late-November (late spring growth) and late February/early March (summer). Rows were scored on 1-10 scale and calibrated by cutting and drying herbage from 15-20 plots. After each assessment, herbage was removed by grazing with ca. 200 sheep for 2-3 days. To compare rows with swards, 33 cultivars and families were evaluated in 4.4 m² sward plots replicated 4 times in a row-column design, sown at 4 kg/ha of viable seed in May 2014. Swards were evaluated by calibrated visual assessment at the same time that the rows were evaluated. Plots were scored on 1-10 scale and ca. 10 quadrats (0.25 m²) cut to ground level for calibration. Both trials received an annual application of 11 kg/ha P as single superphosphate and 60 kg/ha N as urea in winter.

Results

The genotype effect was highly significant ($P < 0.001$) in analyses of variance at all times of observation for both rows and swards. The range among training population families usually exceeded the range for the winter-active cultivars (excluding Australian and Australian II; Table 1). Sirolan was usually the highest yielding cultivar in the main growing season (autumn, winter, spring) and the semi-winter-dormant cultivars, Australian and Australian II, were the lowest (Table 1). Australian and Australian II were the most productive cultivars at the late spring assessment in both years.

Correlations between harvest dates were highest during the autumn-winter-spring period within and between years (Table 2). This is supported by preliminary analyses indicating low family x year interaction at this time of year (not shown). Correlations of late spring growth with growth in the main growing season tended to be lower or non-significant (Table 2).

Correlations between rows and swards for 33 cultivars and families were highest in autumn and winter (~0.8), slightly lower in spring (~0.7), and lower again in late spring and summer (Table 3). Among the 8 cultivars shown in Table 1, correlations over the two years were higher in autumn, winter and spring (0.86-0.97) than in summer ($r = 0.65-0.81$) and late spring ($r = 0.22-0.54$). Scoring was found to be more difficult in late spring and summer when variable amounts of stem were present.

Table 1. Seasonal herbage DM for the highest (TP max) and lowest (TP min) training population families and 8 cultivars in single rows (g/m) and sward plots (kg/ha) at Canberra.

Entry	Summer		Autumn		Winter		Spring (main)		Late spring	
	Row	Sward	Row	Sward	Row	Sward	Row	Sward	Row	Sward
<i>Year 1 - 2015</i>										
TP max	77	-	117	-	93	-	168	-	84	-
TP min	45	-	38	-	39	-	104	-	20	-
Holdfast	62	1458	101	1869	87	1782	166	3214	65	536
Landmaster	61	1507	85	1749	71	1560	146	3214	59	715
Holdfast GT	67	1443	113	1630	98	1520	163	2818	52	511
Advanced AT	64	1219	94	1717	78	1550	151	2667	50	528
Siroso	61	1662	80	2032	89	1819	152	3105	61	731
Sirolan	61	1630	110	2331	100	2191	158	3184	67	803
Australian	51	880	29	842	34	674	83	1985	58	1115
Australian II	44	1108	19	887	33	864	103	2074	59	1255
lsd	11	332	28	313	15	251	19	389	16	290
<i>Year 2 - 2016</i>										
TP max	38	-	55	-	108	-	165	-	171	-
TP min	14	-	30	-	54	-	91	-	45	-
Holdfast	27	546	46	784	91	1985	152	2734	97	1356
Landmaster	27	451	42	755	75	1758	136	2477	93	1465
Holdfast GT	23	412	46	810	90	1990	145	2709	106	1556
Advanced AT	23	390	43	707	75	1726	130	2398	73	1070
Siroso	24	470	43	747	93	2034	149	2658	97	1219
Sirolan	26	562	53	1107	99	2288	144	2837	113	1517
Australian	18	304	36	511	61	1220	104	1800	114	2090
Australian II	19	415	34	685	57	1357	96	1951	76	1663
lsd	7	104	7	173	15	246	20	306	32	480

Table 2. Correlation coefficients between seasonal yields in rows. Values in italics, P<0.05; values in bold, P,0.01.

	S1	A1	W1	Sp1	LSp1	S2	A2	W2	Sp2
Summer 1 (S1)	-								
Autumn (A1)	0.33	-							
Winter 1 (W1)	0.25	0.59	-						
Spring 1 Sp1	0.40	0.47	0.64	-					
Late spring 1 (LSp1)	0.42	0.15	-0.08	0.21	-				
Summer 2 (S2)	0.48	0.22	0.11	0.44	0.64	-			
Autumn 2 (A2)	0.49	0.51	0.47	0.57	0.47	0.64	-		
Winter 2 (W2)	0.23	0.33	0.64	0.51	0.04	0.16	0.61	-	
Spring 2 (Sp2)	0.25	0.24	0.48	0.58	<i>0.13</i>	0.23	0.51	0.76	-
Late spring 2 (LSp2)	0.19	0.07	0.27	0.31	0.31	<i>0.13</i>	0.28	0.47	0.58

Table 3. Correlation coefficients between visually estimated DM yield in rows and swards for 33 entries. Values in italics, P<0.01; values in bold, P<0.001.

Observation period	Year 1	Year 2
Summer	0.62	0.67
Autumn	0.77	0.80
Winter	0.79	0.82
Spring (main)	0.72	0.69
Late spring	0.78	<i>0.53</i>

Discussion

We identified seasonal yield for the most productive medium rainfall zone (550-700 mm) as the highest priority in this project given that breeding for persistence factors has received emphasis for at least 20 years in the CSIRO program. Our training population should address this need and be genetically related enough to a likely breeding population that a relatively small training population will suffice to test GS in phalaris. More summer-dormant cultivars based on germplasm from North Africa and cultivars of the semi-winter dormant (Australian) type were excluded at this stage but can be added to a larger, genetically broader training population, or to new, dedicated populations in the future.

The level of variation among families for herbage yield in the main growing season (autumn, winter, spring), and the relatively high correlations within and between years for growing season herbage yield should be conducive to selection gain in the Canberra environment. However, gain across the intended broader target environment will also depend on G x E interactions. Correlations between rows and swards for a subset of entries were similar to those reported by Smith et al. (2001) for perennial ryegrass families physically harvested or visually estimated. While harvested swards are the ideal for grass evaluation, the much larger number of rows that can be evaluated for a given level of resources makes them a necessary technique in small programs.

We are now in a position to develop relationships between genotypic and phenotypic information based on seasonal yield as well as data on other traits such as seed retention and AI tolerance. Preliminary analyses on early harvests suggest that the correlation between measured and predicted genomic values are similar to those observed in perennial ryegrass (Pembleton et al. 2016).

Acknowledgements

We thank Charlie Veness for technical assistance. The project is funded by MLA Australia.

References

- Baillie RC, Drayton MC, Pembleton LW, Kaur S, Culvenor RA, Smith KF, Spangenberg GC, Forster JW and Cogan NOI (2017). Generation and characterisation of a reference transcriptome for phalaris (*Phalaris aquatica* L.). *Agronomy* 7, 14. (doi.10.3390/agronomy7010014).
- Meuwissen THE, Hayes BJ and Goddard ME (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819-1829.
- Pembleton LW, Baillie RC, Drayton MC, Spangenberg GC, Forster JW, Culvenor RA, Smith KF and Cogan NOI (2016). Development of genomic selection and germplasm discrimination within elite breeding lines of phalaris. In: *Molecular Breeding of Forage and Turf 2016*, Lanzhou University, Lanzhou, China 13-19 August 2016. S1-06.
- Resende RMS, Casler MD and Resende MDV (2014). Genomic selection in forage breeding: accuracy and methods. *Crop Science* 54, 143-156.
- Smith KF, Tasneem M, Kearney GA, Reed KFM and Leonforte A (2001). Evaluation of herbage yield in a forage grass breeding program: comparison of visual rating versus measurement in single-row plots or swards. *Australian Journal of Experimental Agriculture* 41, 1161-1166.