Genetic gain in yield and protein over two cycles of a wheat recurrent selection program

Na Niu1,2, Vivi N. Arief†2, Ian H. DeLacy2, Douglas Lush3, John Sheppard3, Gaisheng Zhang*1 and Mark J. Dieters2

1) Northwest A & F University, Yangling, China
2) The University of Queensland, School of Land, Crop and Food Sciences, Brisbane, Australia
3) Queensland Department of Employment, Economic Department and Innovation, Leslie Research Centre, Toowoomba, Australia

Increasing grain yield and protein are important wheat breeding goals. However, the simultaneous improvement of both traits is complicated by a strong negative correlation and genotype by environment interaction (G×E). A recurrent selection program (RSP) reported here was initiated to produce germplasm, which combined both high yield-potential and increased protein content for the northern wheat-growing region of Australia. Diallel crossing amongst 10 cultivars (6 high yield and 4 high quality cultivars), followed by random mating amongst the F1s were used to produce the founding segregating population. Two years of multi-environment trials were conducted in the GRDC Northern Grain Region (NR) for each of the foundation (RSA0) and first (RSA1) cycles of recurrent selection. Data from the two cycles were analysed to evaluate the genetic gain achieved over the two cycles of selection, based on five potential selection strategies. Selection based on yield or protein or yield-adjusted protein only increased selected trait but reduced the other traits. However, application of selection based on yield and protein or yield and yield-adjusted protein simultaneously increased both yield (4.7%, 6.0% for RSA0 and 4.0%, 4.2% for RSA1 respectively) and protein (2.1%, 1.3% for RSA0 and 1.7%, 1.6% for RSA1 respectively).

Key Words: wheat, genetic gain, recurrent selection, principal component analysis, grain yield, grain protein.

Introduction

Grain protein is an important factor affecting grain quality and thus a key determinant of both end use and market value in hexaploid wheat. However, it is known that grain protein is negatively correlated with grain yield in wheat (Bhatia 1975, Mc Neal et al. 1982, O'Brien and Ronal ds 1984, Peterson et al. 1992, Liu et al. 1995, Boggini et al. 1997), hence selection to increase one trait will tend to decrease the second trait (Jenner et al. 1991, Rharrabti et al. 2001). Despite this negative correlation, several studies have reported simultaneous improvement of grain yield and grain protein (Torp 1979, Cox et al. 1985, Noaman et al. 1990).

Recurrent selection is a common breeding method that plant breeders use for cultivar improvement, which is based on successive cycles of selection, testing, and recombination to accumulate favourable alleles of genes controlling a quantitative character, such as yield, resistance, protein content in a population without a significant loss of genetic variability (Olmedo-Arcega et al. 1995, Beeck et al. 2008). Due to the negative correlation between grain yield and protein, recurrent selection should be a more effective approach to the joint improvement of yield and protein (Delzer et al. 1995, Wilcox 1998, Frey and Holland 1999, Singh et al. 1999, Wiersma et al. 2001). A recurrent selection program (RSP) was initiated to produce germplasm that combines both high-yield and high-protein potential adapted to the northern wheat-growing regions of Australia. In this paper, we reported genetic gain obtained from two cycles of the RSP for both grain yield and grain protein.

Materials and Methods

Experimental design

The RSP reported here was initiated in 1996 to jointly improve both high grain yield and protein concentration. The base population was initiated by diallel crossing amongst 10 cultivars (6 high yield cultivars and 4 good quality cultivars) with other good agronomic performance (e.g. multiple disease resistances). This population is referred as RSA. In the foundation cycle (i.e. RSA0), the ten parental cultivars were inter-mated in a half-dialled mating design, and each of 45 different F1’s produced were randomly mated to produce a S0 population of approximately 10000 plants. In the foundation cycle all grains produced from random mating were bulked, but in the subsequent cycle grains from each female parent were kept separate in the S0 generation. Single plants (approximately 2000 in RSA0 and 750 in...
RSA1) were selected in the S₀ population on the basis of disease resistance, height, maturity and lodging, and S₁ grain harvested separately from each selected plant. S₁ plants were then grown over summer, to produce grain for establishment of each S₁ family in multi-environment trials during the subsequent two winters. At the end of the foundation cycle, the top 20 S₁ families were selected as the parents for the next cycle and randomly mated as described above. In order to understand the practical method, the procedure of the second cycle of RSA (i.e. RSA1) was shown in Fig. 1.

In foundation cycle (i.e. RSA0) 772 families were tested at five locations (Narrabri, Central Queensland, Billa Billa, Gatton and Jimbour) in 1998 and repeated in 1999, while in the subsequent cycle (i.e. RSA1), for the first year of testing (2007) 737 families were tested in six locations (Biloela, Coonamble, Myall Valley, Orion, Weemelah), while for the second year (2008) of testing a subset of 181 families were tested in five locations (Billa Billa, Biloela, Culara, Gatton, and Tulloona). All locations were representative of the wheat growing area in Queensland and northern New South Wales. Note that the lag between these two cycles resulted from funding issues rather than any biological or genetic constraints. Each cycle can now be completed in 3 years.

For analysis, each location and year was treated as a different trial. Data were collected for grain yield and protein concentration on a plot mean basis. Near infrared (NIR) spectra were collected using MPA (Multi Purpose FT-NIR Analyser, Bruker Corporation) with the OPUS 6.0 software. Protein concentrations were estimated based on NIR spectroscopy (Osborne et al. 1993) and expressed as percentage. For RSA0, protein data were only available from five trials (out of ten trials across two years) and for RSA1 from ten trials (out of eleven trials across two years).

Data analysis

Two-stage mixed model analyses (Cullies et al. 1996, Smith et al. 2001) were performed for grain yield and protein. Spatial model (Cullis and Gleeson 1998, Cullis et al. 1998) was fitted in the first stage for each trait at each site. The BLUE (Best linear unbiased estimators) and weights (Cullis et al. 1996) were calculated and used in the second stage analysis. Protein level is confounded with yield level due to the negative phenotypic correlation between yield and protein level (believed to be largely a dilution effect). To obtain a better estimate of the intrinsic protein potential of each family, yield was used as covariate in the one-stage spatial mixed model analysis (Smith et al. 2005) to estimate what we have termed yield-adjusted protein (a covariate could not be fitted in two-stage analysis). All analyses were conducted using mixed model approach based on Restricted Maximum Likelihood method (Patterson and Thompson 1971) implemented in ASREML (Gilmour et al. 2001).

The BLUPs (Best Linear Unbiased Predictors) for each family, trait and trial from the second-stage analyses were used to performed pattern analysis (Williams 1976, Cooper and DeLacy 1994). For clustering squared Euclidean distance was used as the distance measure and Incremental Sum of Squares (Ward 1963) was used as the clustering strategy (Cooper and DeLacy 1994). For ordination singular value decomposition was used. The results of pattern analysis were displayed in a series of biplots (Gabriel 1971) using symmetrical scaling (Kroonenberg 1997).

Selection

Five potential selection strategies were investigated to identify 20 S₁ families to be used as parents of the next cycle: top-20 families selected based on (1) yield only; (2) protein only; and (3) yield-adjusted protein only; (4) joint using of yield and protein (5) joint using of yield and yield-adjusted protein. Response to selection for these 20-selected families were calculated by comparing mean (for yield and protein) to their population from which they were selected (i.e. either RSA0 or RSA1) and expressed as the percentage difference.

Results

Yield-adjusted protein

Scatter plots for yield, protein and adjusted protein from RSA0 and RSA1 (Fig. 2) indicate a stronger negative correlation in RSA0 than in RSA1 between yield and protein (compare Fig. 2A and 2D). Correlation between yield and yield-adjusted protein, although still negative, was increased compared with the correlation between yield and protein.
Genetic gain in yield and protein over wheat recurrent selection

The first two principal components explained 56.6% of the variability in RSA0 and 82.7% in RSA1. For both RSA0 and RSA1, the negative correlation between yield and protein, the less negative correlation between yield and yield-adjusted protein still could be seen in the biplots (Fig. 3 and Fig. 4), as shown by the angle between vectors for yield and protein. The positive correlation between protein and yield-adjusted protein could also be seen (Fig. 3 and Fig. 4). Although in this case protein and adjusted protein are highly correlated, adjusted protein tended to decrease negative correlated with yield. Thus selection based on adjusted protein would select lines with higher protein level regardless the yield level.

Depending on the selection strategies, the 20-selected families occupied a different space in the biplots (Fig. 3 and Fig. 4): those selected on only tended to occupy the space closer to yield vectors; those selected on protein and yield-adjusted protein respectively tended to occupy the space closer to the protein and yield-adjusted protein vectors. However, families selected jointly using yield and yield-adjusted protein, occupied the space between yield and yield-adjusted protein vectors, with RSA1 tending to have more families in this space than RSA0.

Selection

The first two principal components explained 56.6% of the variability in RSA0 and 82.7% in RSA1. For both RSA0 and RSA1, the negative correlation between yield and protein, the less negative correlation between yield and yield-adjusted protein still could be seen in the biplots (Fig. 3 and Fig. 4), as shown by the angle between vectors for yield and protein. The positive correlation between protein and yield-adjusted protein could also be seen (Fig. 3 and Fig. 4). Although in this case protein and adjusted protein are highly correlated, adjusted protein tended to decrease negative correlated with yield. Thus selection based on adjusted protein would select lines with higher protein level regardless the yield level.

Depending on the selection strategies, the 20-selected families occupied a different space in the biplots (Fig. 3 and Fig. 4): those selected on yield only tended to occupy the space closer to yield vectors; those selected on protein and yield-adjusted protein respectively tended to occupy the space closer to the protein and yield-adjusted protein vectors. However, families selected jointly using yield and yield-adjusted protein, occupied the space between yield and yield-adjusted protein vectors, with RSA1 tending to have more families in this space than RSA0.

Response to selection

As expected, selection based only on yield maximised gain in yield but also led to a reduction in protein and vice versa for selection only on protein and yield-adjusted protein (Table 1). Selection based on combined yield and protein not only increased the response for yield but also for protein in both RSA0 and RSA1. The similar result for the selection based on combined yield and yield-adjusted protein due to the very strong correlation between protein and yield-adjusted protein. Selection based on either model identified many families in common. However, the response for yield increased significantly by the yield and yield-adjusted protein model as compared to the yield and protein model in RSA0, but also resulted in a very small decrease for protein. In RSA1, there are no significant differences between the two joint selection models for two traits. Compared RSA0 to RSA1, improvement in yield was higher in RSA0 than in RSA1 based on single selection, but similar for joint selection. For protein both RSA0 and RSA1 selected families showed similar levels of improvement (Table 1).
**Discussion**

Improving grain yield and protein content is the important problem confronting agriculture. Increasing of yield has been achieved by developing cultivars with poor protein content (Jenner et al. 1991, Rharrabti et al. 2001) because of negative correlation between yield and protein content (Bhatia 1975, McNeal et al. 1982, O’Brien and Ronalds 1984, Peterson et al. 1992, Liu et al. 1995, Boggini et al. 1997). In our research, negative correlation was confirmed to exist between yield and protein content. Selection one trait would decrease the other trait if just using single selection strategy. However, it is entirely possible to increase of grain yield and protein simultaneously. This result was identical with some researches (Torp 1979, Cox et al. 1985, Noaman et al. 1990). Our analyses showed that recurrent selection program can be used to simultaneously improve both yield and protein despite the strong negative genetic correlation between these two traits in hexaploid wheat. Grain yield and protein could be increased simultaneously by means of joint using two traits. After one cycle of recurrent selection, the negative correlation between yield and protein appears to have decreased.

Many researchers have focused on the evaluation of recurrent selection programs in cereals (Feil 1992, Sláfer et al. 1994). Several authors have applied recurrent selection successfully for improving traits like grain proteins (Delzer et al. 1995), grain weight (Busch and Kofoid 1982) and yield (Bravant et al. 1991, Olmedo-Arceaga et al. 1995, Dubois et al. 1998). However, most of breeding programs concerned only one trait. In this study, grain yield and protein content were covered simultaneously, and the genetic gains for two cycles were evaluated. Since selection strategy is important to breeding objective, we compare five strategies through calculate the responses to selection. Our results showed that if selection was performed only based on one trait, the largest response was gotten, and the least response was gotten for unselected trait, since the negative correlation existed between grain yield and protein. The objective of RSA program is to improve the grain yield and maintain the protein content. So we should carry selection based on the joint information. From our study, joint selection for both traits simultaneously is able to improve grain yield while also increasing the level of grain protein in both cycles (Table 1).
The use of yield-adjusted protein also enabled selection for families with higher intrinsic protein, and tended to reduce the strength of the negative correlation between yield and protein. However, in selection, yield and yield-adjusted protein could increase the responses for yield significantly and decrease protein slightly in RSA0. There are not significant differences between the two joint selection models for two traits in RSA1. Thus, the yield and yield-adjusted protein model may play a more role important for selection in the early cycle. Compared RSA0 to RSA1, improvement in yield was higher in RSA0 than in RSA1 based on single selection, but similar for joint selection, because the yield trait is more affected by environment. Some of top 20 lines selected in RSA1 were affected by rust diseases and lodging before harvesting, which would probably make response decrease sharply.

The pattern analysis showed that, in order to obtained families that had both, high-yield and high-protein, the families should occupy the space between yield and protein vectors (Fig. 3 and Fig. 4). After two-cycle of selections a greater proportion of the families selected in RSA1 occupied this space (Fig. 3 and Fig. 4), and it is anticipated that future

**Table 1.** Response to selection (%) for 20 selected families in RSA0 and RSA1 families based on five selection strategies

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Selection strategy</th>
<th>Yield (%)</th>
<th>Significance of difference</th>
<th>Protein (%)</th>
<th>Significance of difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSA0</td>
<td>YLD</td>
<td>12.0</td>
<td>a</td>
<td>−3.0</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>PRT</td>
<td>−6.1</td>
<td>e</td>
<td>5.1</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>YAP</td>
<td>−3.1</td>
<td>d</td>
<td>4.7</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>YLD+PRT</td>
<td>4.7</td>
<td>c</td>
<td>2.1</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>YLD+YAP</td>
<td>6.0</td>
<td>b</td>
<td>1.3</td>
<td>b</td>
</tr>
<tr>
<td>RSA1</td>
<td>YLD</td>
<td>6.3</td>
<td>a</td>
<td>−2.7</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>PRT</td>
<td>−3.1</td>
<td>c</td>
<td>5.1</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>YAP</td>
<td>−1.7</td>
<td>c</td>
<td>4.8</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>YLD+PRT</td>
<td>4.0</td>
<td>b</td>
<td>1.7</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>YLD+YAP</td>
<td>4.2</td>
<td>b</td>
<td>1.6</td>
<td>b</td>
</tr>
</tbody>
</table>

YLD for selection based on yield; PRT for selection based on protein; YAP for selection based on yield-adjusted protein; YLD+PRT for based on yield and protein; and YLD+YAP for selection based on yield and yield-adjusted protein. Percentage was calculated as the ratio of mean difference between selected lines and the population with population mean. Positive sign indicated increased in the mean of selected lines and vice versa for negative sign. Different letters in the same column represent significant differences at P = 0.05 by Ducan’s multiple range test.
cycles of recurrent selection will increasingly move the selected genotypes into this trait-space.

**Acknowledgements**

We thank the Grains Research and Development Corporation for funding the recurrent selection program. We also thank the team of people at the Leslie Research Centre of Queensland Department of Employment, Economic Development and Innovation for conducting this program and providing with data. We thank the supporting from the National High Technology Research and Development Program (“863” Program) of China (2009AA101102). We also thank the TANG Foundation for the research of breeding from North West A & F University (A212020909).

**Literature Cited**


