

Supplementary information: Details of statistical analyses

1. Analysis of individual subsoil constraint traits

In the experiments used to characterise the 52 indicator genotypes the model considered was

$$\mathbf{y} = \mathbf{X}\boldsymbol{\tau} + \mathbf{Z}\mathbf{u} + \mathbf{Z}_g\mathbf{g} + \mathbf{e}$$

where \mathbf{y} is the $n \times 1$ vector of trait observations, the term $\mathbf{X}\boldsymbol{\tau}$ represents the fixed effects part of the model, the term $\mathbf{Z}\mathbf{u}$ represents the random effects given by the structure of the experiment and term $\mathbf{Z}_g\mathbf{g}$ represents the variety effects. The \mathbf{e} term represents the residual error and is assumed to be independent normally distributed with constant variance. The terms in the model vary depending on the number of treatments in the experiment.

Single treatment assays. The assays for Al tolerance, RLN resistance, CID and seminal root angle involved one treatment only. In these experiments, the fixed effect consists of a mean while the random effects involve possible blocking or Replicate effects. The variety effects, \mathbf{g} , are assumed to have zero mean and constant variance. The best linear unbiased predictions (BLUPs) of the 52 indicator variety effects were calculated after analysis.

Two treatment assays. For experiments with more than one treatment (e.g., B tolerance, Zn efficiency) the non-genetic fixed effects consist of a mean and a Treatment effect whereas the non-genetic random effects typically consist of a Replicate effect and Treatment by Replicate effect and are assumed to be independently normally distributed with constant variances. The

error is assumed to be independent normally distributed with zero mean and variance depending on the Treatment group or level.

For 2 treatments, the term $\mathbf{Z}_g \mathbf{g}$ represents a treatment by Variety interaction, where the matrix \mathbf{Z}_g assigns the appropriate Variety to each observation and \mathbf{g} is a size $2g \times 1$ vector of random effects with distribution

$$\begin{pmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \end{pmatrix} \sim N \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{pmatrix} \otimes I_g \right]$$

This structure therefore provides variety variances for both treatment levels as well as a covariance between treatment levels. This is a sensible model for multi-treatment trials designed to provide phenotypes. The linear mixed model provides BLUPs for each variety corresponding to each treatment level.

The varietal efficiency was assessed by examining the performance of a variety under the treatment relative to its performance in the Nil treatment using regression analysis. Focussing on the two treatment case, for example with B tolerance, the efficiency the varieties under the Boron treatment given the Nil treatment can be visualized as a regression with the slope reflecting the rate of change of the Boron treatment relative to the Nil treatment. Deviations around the regression will indicate the relative performance of a variety or its efficiency. Because \mathbf{g} is a random effect, the regression can be formed simply by considering the conditional distribution of \mathbf{g}_1 (Boron) given \mathbf{g}_0 (Nil), namely

$$\mathbf{g}_e = \mathbf{g}_1 | \mathbf{g}_0 \sim N(E(\mathbf{g}_1 - \sigma_{10}(\sigma_0^2)^{-1} \mathbf{g}_0), \sigma_1^2 - \sigma_{10}(\sigma_0^2)^{-1} \sigma_{01}) \quad (1.1)$$

where $E()$ is an expectation. The expression $\sigma_{10}(\sigma_0^2)^{-1}$ becomes a regression coefficient for the regression of \mathbf{g}_1 on \mathbf{g}_0 and the efficiency score is therefore a residual where varieties with

large residuals represent lines that are either efficient or inefficient for Boron tolerance. The efficiency scores are empirically calculated using

$$\tilde{\mathbf{g}}_e = \tilde{\mathbf{g}}_1 - \hat{\sigma}_{10}(\hat{\sigma}_0^2)^{-1}\tilde{\mathbf{g}}_0$$

where $\tilde{\mathbf{g}}_1$ and $\tilde{\mathbf{g}}_0$ are the BLUPs for the Boron and Nil treatment variety effects and the regression slope contains substituted REML variance parameter estimates. Similar calculations were done for all experiments containing two treatments.

2. Genotype x Environment analysis

Let $\mathbf{y}_j = (y_{1j}, \dots, y_{n_jj})$ be the yield response for the j th trial ($j = 1, \dots, 233$) and n_j the trial size. The initial linear mixed model for the j th trial is

$$\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{Z}_j\mathbf{u}_j + \mathbf{e}_j$$

where $\boldsymbol{\beta}_j$ and \mathbf{u}_j represent the fixed and random effects respectively and \mathbf{X}_j and \mathbf{Z}_j are their associated design matrices. The fixed component of the model is partitioned to give

$$\mathbf{X}_j\boldsymbol{\beta}_j = \mathbf{X}_j^{(g)}\boldsymbol{\beta}_j^{(g)} + \mathbf{X}_j^{(a)}\boldsymbol{\beta}_j^{(a)}$$

where $\boldsymbol{\beta}_j^{(g)} = (\beta_{1j}^{(g)}, \dots, \beta_{g_jj}^{(g)})$ represent the vector of Genotype or variety means ($1, \dots, g_j$) present in the j th trial. The remaining effects, $\boldsymbol{\beta}_j^{(a)}$, are used to estimate covariate or possible spatial information that might be present for that particular trial.

The vectors \mathbf{u}_j and \mathbf{e}_j are considered independent with distributions $\mathbf{u}_j \sim N(\mathbf{0}, \sigma^2 \mathbf{G}_j)$ and $\mathbf{e}_j \sim N(\mathbf{0}, \sigma_j^2 \boldsymbol{\Sigma}_{jr} \otimes \boldsymbol{\Sigma}_{jc})$ respectively. The random effects \mathbf{u}_j are considered to be design effects associated with the j th trial which may include possible Replicate effects, simple spatial Row or Column effects and, if required, non-linear spatial effects that may be present

in the field. Thus, the structure of \mathbf{G}_j may vary from trial to trial. The error component of the model, \mathbf{e}_j is a two-dimensional spatial trend process with a separable variance structure. For field trials it is generally accepted that this structure consist of a separable $\text{AR}(1) \times \text{AR}(1)$ process and therefore Σ_{jr} and Σ_{jc} represent correlation matrices for Row and Columns respectively which each contain a parameterisation for an autoregressive process of order 1 in the their respective direction. The parameter, σ_j^2 represents the residual spatial variation or plot error associated with the j th trial.

The data for the second stage is formed by combining the estimated Genotype means for the 52 genotypes for the individual trials from the first stage, namely

$$\mathbf{y}^* = [(\hat{\beta}_{11}^{(g)}, \dots, \hat{\beta}_{g_{11}}^{(g)}), \dots, (\hat{\beta}_{1j}^{(g)}, \dots, \hat{\beta}_{g_{jj}}^{(g)}), \dots, (\hat{\beta}_{1(233)}^{(g)}, \dots, \hat{\beta}_{g_{233}(233)}^{(g)})]$$

The second stage Factor Analytic model is now given by

$$\mathbf{y}^* = \mathbf{X}^* \boldsymbol{\tau} + \mathbf{Z}^* \mathbf{g}_v + \mathbf{e}^* \quad (2.1)$$

where, again, $\boldsymbol{\tau}$ and \mathbf{g}_v are fixed and random effects and \mathbf{X}^* and \mathbf{Z}^* are their associated design matrices. For this particular model the fixed effect represents the vector of mean Environment effects, $\boldsymbol{\tau} = (\tau_1, \dots, \tau_{233})$. In this model \mathbf{u}_g is a $(52 \times 233) \times 1$ vector of genetic effects for the 52 genotypes in the 233 environments. These effects are assumed to have an underlying Factor Analytic model of the form

$$\mathbf{g}_v = \sum_{i=1}^k (\boldsymbol{\lambda}_i \otimes \mathbf{I}_g) \mathbf{f}_i + \boldsymbol{\delta} = (\boldsymbol{\Lambda} \otimes \mathbf{I}_g) \mathbf{f} + \boldsymbol{\delta}$$

where \mathbf{f}_i is the score for the i th hypothetical factor across all varieties, $\boldsymbol{\lambda}_i$ is the associated loading across all environments and $\boldsymbol{\delta} = (\delta_1, \dots, \delta_{233})$ are independent residuals with

$\delta_j \sim N(\mathbf{0}, \psi_{vj} \mathbf{I})$ for the j th environment. Therefore the variance matrix for the Indicator varieties in each Environment can be expressed as

$$\text{var}(\mathbf{g}_v) = \mathbf{G}_e \otimes \mathbf{I}_g$$

where $\mathbf{G}_e = (\mathbf{A}\mathbf{A}^T + \mathbf{\Psi}_v)$ is the genetic variance/covariance matrix for all Environments.

The joint distribution of $(\mathbf{u}_g, \mathbf{e}^*)$ has the form

$$\begin{pmatrix} \mathbf{g}_v \\ \mathbf{e}^* \end{pmatrix} \sim N \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} (\mathbf{A}\mathbf{A}^T + \mathbf{\Psi}_v) \otimes \mathbf{I}_g & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix} \right]$$

where \mathbf{R} is a block diagonal matrix with blocks \mathbf{R}_j , $j = 1, \dots, 233$. This block diagonal matrix is the conditional variance of the new response, \mathbf{y}^* given \mathbf{g}_v and represents the uncertainty due to the data being mean estimates. In general, the block diagonal matrices \mathbf{R}_j , $j = 1, \dots, 234$ are replaced by known estimates $\tilde{\mathbf{R}}_j$, $j = 1, \dots, 233$ obtained from the analysis of the individual trials. For the j th trial this is calculated using

$$\tilde{\mathbf{R}}_j = \text{var}(\mathbf{y}_j^*) = \tilde{\sigma}_j^2 (\mathbf{X}_j^{(g)T} \tilde{\mathbf{H}}_j^{-1} \mathbf{X}_j^{(g)})^{-1}$$

where $\tilde{\mathbf{H}}_j = \tilde{\sigma}_j^2 (\mathbf{Z}_j \tilde{\mathbf{G}}_j \mathbf{Z}_j^T + \tilde{\mathbf{\Sigma}}_{jr} \otimes \tilde{\mathbf{\Sigma}}_{jc})$. It is computationally complex to use the full block diagonal matrices in the second stage of the analysis (see Frensham et. al., 1997). A standard approximate approach to accommodate for known heterogeneity in linear models is to consider the inverse of the variances of the data points as weights. However, as $\tilde{\mathbf{R}}_j$, $j = 1, \dots, 233$ are not diagonal matrices the most appropriate weights are given by $\mathbf{w}_j = \text{diag}[\tilde{\mathbf{R}}_j^{-1}]$, $j = 1, \dots, 233$. Therefore each Genotype mean for the second stage has its own unique weight that is incorporated into the second stage of the analysis.

3. Analysis of contributions of subsoil constraint traits to yield

For this analysis to proceed a simpler model was proposed of the form

$$\mathbf{y}^* = \mathbf{X}^* \boldsymbol{\tau} + \sum_{i=1}^{10} \mathbf{Z}_i \mathbf{g}_i + \mathbf{Z}^* \mathbf{g}_v + \mathbf{e}^*$$

where $\mathbf{X}^* \boldsymbol{\tau}$ and \mathbf{e}^* are the fixed and error components of the model from (2.1). \mathbf{g}_i are a vector of random coefficients across all environments for the i th subsoil constraint trait with distribution $\mathbf{g}_i \sim N(\mathbf{0}, \boldsymbol{\Psi}_i \otimes \mathbf{I}_g)$ with $\boldsymbol{\Psi}_i = \text{diag}(\psi_{i1}, \dots, \psi_{i233})$. In this model $\mathbf{Z}_i = (\tilde{\mathbf{g}}_i \otimes \mathbf{1}_g)$ is a covariate where $\tilde{\mathbf{g}}_i$ are the genetic BLUPs or efficiency scores of the 52 genotypes for the i th subsoil constraint trait obtained from the analyses conducted in Appendix 1. Note, that each of the predictions $(\tilde{\mathbf{g}}_1, \dots, \tilde{\mathbf{g}}_{10})$ incorporated into the meta-analysis have their own prediction error. This induces an additional complex ‘error-in-variables’ structure to the model which, for simplification purposes, is ignored in the analysis that follows. Similar to (2.1), \mathbf{g}_v represent the variety effects except that the complexity of its distribution is reduced to $\mathbf{g}_v \sim N(\mathbf{0}, \boldsymbol{\Psi}_v \otimes \mathbf{I}_g)$ with $\boldsymbol{\Psi}_v = \text{diag}(\psi_{v1}, \dots, \psi_{v233})$.

With this simplification, at any particular site, the genetic contribution of an individual subsoil constraint trait can be summarised by considering the estimated genetic variance of the individual subsoil constraint trait proportional to the total estimated genetic variance. For the i th subsoil constraint trait at site j this is

$$\frac{\hat{\psi}_{ij}}{(\sum_{i=1}^{10} \hat{\psi}_{ij}) + \hat{\psi}_{vj}}$$

**Supplementary Table 1. Summary of the distributions in relative yield
across 233 sites of varieties showing greater tolerance to subsoil constraint**

traits

Subsoil constraint trait	Relative grain yield (%)				
	Minimum	Quartile	Median	Quartile	Maximum
		1		3	
Al tolerance	72.1	100.3	105.3	109.4	169.4
Boron tolerance	71.7	101.7	106.4	112.1	275.1
High pH tolerance	39.2	95.9	101.1	104.2	148.7
High Mn concentration	41.6	99.3	102.8	107.0	153.1
Mn efficiency	73.4	95.6	102.9	110.1	252.4
Low Na concentration	79.0	101.4	107.4	116.5	306.5
Na relative dry matter	62.6	93.8	97.9	101.9	130.3
Zn efficiency	64.0	94.9	98.2	101.5	130.1
<i>P. neglectus</i> resistance	61.3	95.6	100.1	104.9	134.5
Root penetration	75.7	94.2	98.4	103.1	199.2
Narrow root angle	78.2	99.6	102.7	106.7	176.9
Low CID	85.4	100.0	101.6	103.6	119.7
Early flowering	80.3	99.8	104.8	109.6	127.9

Supplementary Table 2. Estimation of the phenotypic effects of individual subsoil constraint traits on the yield of wheat in different rainfall quartiles in Western Australia, South Australia and Victoria. The analysis is based on trials in which the subsoil constraint trait made a significant contribution to the genetic variation in yield and the rainfall quartiles are based on all sites in used in the analysis.

The data compare the yield of the 10 best and 10 worse varieties for each trait.

Trait	Variety rating	Rainfall quartile			
		1 st quartile	2 nd quartile	3 rd quartile	4 th quartile
		(t/ha)			
Al tolerance	Intolerant	1.64	2.13	2.54	3.32
	Tolerant	1.82 (+11.1%)	2.36 (+10.9%)	2.78 (+9.3%)	3.56 (+7.2%)
	Signif.	P<0.001	P<0.001	P<0.001	P<0.01
	df	12	15	11	12
B tolerance	Intolerant	1.14	1.67	1.87	2.87
	Tolerant	1.35 (+17.9%)	1.90 (+14.0%)	2.18 (+16.1%)	3.30 (+15.2%)
	Signif.	P<0.001	P<0.001	P<0.001	P<0.01
	df	11	11	13	6
High pH tol.	Intolerant	1.14	2.01	2.24	4.94
	Tolerant	1.19 (+4.9%)	2.09 (+4.4%)	2.31 (+3.1%)	5.06 (+2.6)
	Signif.	P<0.05	P<0.01	NS	NS
	df	8	8	5	4
Mn conc	Low	1.25	2.09	2.62	2.91
	High	1.38 (+10.1%)	2.35 (+12.1%)	2.94 (+12.3%)	3.13 (+7.2%)
	Signif.	P<0.005	P<0.01	P<0.05	P<0.01
	df	6	5	8	8
Mn efficiency	Inefficient	2.32	2.30	2.25	3.72
	Efficient	2.58 (+10.8%)	2.52 (+9.6%)	2.49 (+8.3%)	3.91 (+5.4)
	Signif.	- ^A	P<0.01	P<0.01	P<0.05
	df		4	9	7
Na conc	High	1.37	1.61	2.23	3.20
	Low	1.57 (+14.2%)	1.88 (+17.1%)	2.64 (+18.4%)	3.53 (+10.1%)
	Signif.	P<0.001	P<0.001	P<0.01	P<0.001
	df	26	24	13	10
Na relative DM	Low	1.87	2.23	2.10	4.28
	High	1.89	2.26	2.27 (+8.0%)	4.58 (+7.0%)
	Signif.	NS	NS	P<0.10	P<0.05
	df	3	7	8	5
Zn efficiency	Inefficient	0.77	1.58	3.29	5.73

	Efficient	0.82 (+6.5%)	1.61 (+1.9%)	3.38 (+2.6%)	5.78
	Signif.	P<0.05	NS	- ^A	NS
	df	5	6	-	1
<i>P. neglectus</i> resist	Susceptible	1.49	1.87	2.21	3.14
	Resistant	1.57 (+5.0%)	1.99 (+6.5%)	2.39 (+8.0%)	3.41 (+8.5%)
	Signif.	P<0.001	P<0.001	P<0.001	P<0.001
	df	15	15	20	14
Relative root growth	Low	1.47	1.44	2.57	4.14
	High	1.37 (-6.6%)	1.54 (-6.9%)	2.37 (-7.8%)	3.91 (-5.4%)
	Signif.	P<0.005	P<0.01	P<0.001	P<0.005
	df	12	9	16	13
Root angle	Wide	1.36	1.96	2.45	3.15
	Narrow	1.45(+7.1%)	2.18(+11.3%)	2.69 (+9.5%)	3.55 (+12.7%)
	Signif.	P<0.001	P<0.01	P<0.001	P<0.001
	df	12	14	9	17
Low CID	High	1.48	2.20	1.99	4.16
	Low	1.53 (+3.6%)	2.32 (+5.9%)	2.04 (+2.5%)	4.31 (+3.5%)
	Signif.	P<0.10	P<0.01	P<0.05	P<0.01
	df	4	7	6	7
High CID	Low	1.60	1.73	2.04	2.68
	High	1.61	1.77	1.98	2.72
	Signif.	NS	NS	NS	NS
	df	2	4	1	1
Early maturity	Late	1.26	1.83	2.12	3.27
	Early	1.50 (+19.1%)	2.17 (+18.4%)	2.58 (+21.2%)	3.72 (+14.1%)
	Signif	P<0.001	P<0.001	P<0.001	P<0.001
	df	17	15	14	14
Late maturity	Early	1.87	1.82	2.85	3.20
	Late	2.09 (+11.9%)	2.14(+17.4%)	2.83	4.06 (+26.7%)
	Signif	P<0.05	P<0.05	NS	P<0.05
	df	3	7	1	7

^AOnly one site had a significant relationship between tolerance and yield.

Supplementary Fig 1. Variation among 52 genotypes of bread wheat for the subsoil constrain traits used to characterise each variety. For Al tolerance, C isotope discrimination, shoot Mn concentration, *P. neglectus* numbers, seminal root angle and tissue Na⁺ concentration, for which a single treatment was analysed, values are BLUPs for each variety. For the remaining traits, each variety was exposed to two treatments (low and high B, low and high pH, low and high Mn, low and high soil strength) and the values are the deviations from the relationship between the BLUPs of the low treatment vs BLUPs of the high treatment. A negative deviation represents lower than average efficiency and a positive deviation represents a higher than average efficiency. Data for *P. neglectus* numbers were transformed to log (number/1000) before analysis and the data are shown as the BLUPS of the transformed data.





