Text S2. Application of expVIP to rice allows integration of previous studies

The rapid progress in sequencing technologies has meant that genomes and transcriptome references for a species are constantly improved. In some species, multiple references are available, making it difficult to compare results between studies. This problem is exemplified in rice, where two different genome annotations are widely used: the Rice Annotation Project (RAP) and Michigan State University (MSU) gene models (Ohyanagi et al., 2006; Ouyang et al., 2007).

To test the use of expVIP to integrate data we analysed two studies (SRA: DRP000716 and SRP028766) which examined gene expression changes in rice in response to phosphate starvation. DRP000716 compared the response to 22 days of phosphate (Pi) starvation in four rice varieties with varying levels of tolerance to Pi stress: the *japonica* cultivar Nipponbare with low tolerance, two *japonica* cultivars with higher tolerance IAC 25 and Vary Lava, and the *indica* cultivar Kasalath known to be highly tolerant to Pi stress (Oono et al., 2013). The second study SRP028766 investigated how the *japonica* cultivar Nipponbare responded to a time-course of Pi starvation at 1 h, 6 h, 24 h, 3 days, 7 days and 21 days (Secco et al., 2013). Both studies used 2 week old seedlings grown in hydroponic conditions which enables their comparison, however each study used a different annotation of the rice genome (DRP000716 used RAP, SRP028766 used MSU). We used expVIP to align and quantify gene expression for both studies, using the RAP gene models (release date 31.03.2015, from http://rapdb.dna.affrc.go.jp) as a common reference. We used sleuth to identify genes which were differentially expressed between varieties in DRP000716 (each variety under phosphate starvation was compared to the same variety without phosphate starvation) and between time-points in SRP028766 (phosphate starved plants were compared to plants at the same time-point which were in phosphate sufficient conditions).

Validation of previous results

DRP000716 – Inter-variety comparison**:** We found that similar numbers of genes were up and downregulated in all four varieties (Figure S3A), although Kasalath roots had fewer genes downregulated than in any other variety. We identified approximately 2-fold fewer

differentially expressed transcripts than published previously (Oono et al., 2013), which is likely due to the use of a defined reference in our study, whereas before *de novo* transcripts including multiple isoforms were also assembled. We found a total of 163 genes upregulated and 34 genes downregulated in roots and shoots of all varieties. Amongst these upregulated genes, many are known to be involved in phosphate response including *SPX1* and *SPX3* (Wang et al., 2009) and the phosphate transporter *PHT1;4* (Zhang et al., 2015). The downregulated genes included genes related to primary metabolism, e.g. ribulose bisphosphate carboxylase, and genes involved in abiotic stress responses such as *RISBZ5*, a potential negative regulator of drought and cold stress response (Liu et al., 2012).

SRP028766 – time-course**:** As previously reported we found that relatively few genes were induced within a short time period after imposition of phosphorous starvation (1h, 6h and 24 h; Figure S3B). At 3 days many genes become differentially expressed in roots, whereas there are still few genes differentially expressed in shoots. At 7 and 21 days several thousand genes are up and downregulated in roots and shoots. These results correspond well to the previously reported trends (Secco et al., 2013): we found that the early response (1 h - 3 days) to phosphate starvation involves suites of different genes at each time-point in roots and as previously reported no genes were differentially expressed in roots in common between all early time-points.

Comparison between DRP000716 and SRP028766

First we investigated whether the two different studies identified similar genes to be differentially expressed in the cultivar Nipponbare after 21 or 22 days under phosphate starvation. We found that in total 1,565 and 2,001 genes were differentially expressed in shoots and roots, respectively, across both studies (Figure S3C). We identified fewer differentially expressed genes in DRP000716 than in SRP028766, which may reflect the lower number of reads mapped in the former (15.5 million and 46.2 million, respectively). We found that amongst genes which were differentially expressed in both studies, there was a higher correlation of fold change in genes expressed in the roots (R^2 = 0.64) than in genes expressed in the shoots (R^2= 0.38) (Figure S3D). This suggests that changes in root gene

expression were more consistent between studies (more shared genes, and more highly correlated fold changes in expression), and for this reason we focused our analysis on genes differentially expressed in roots.

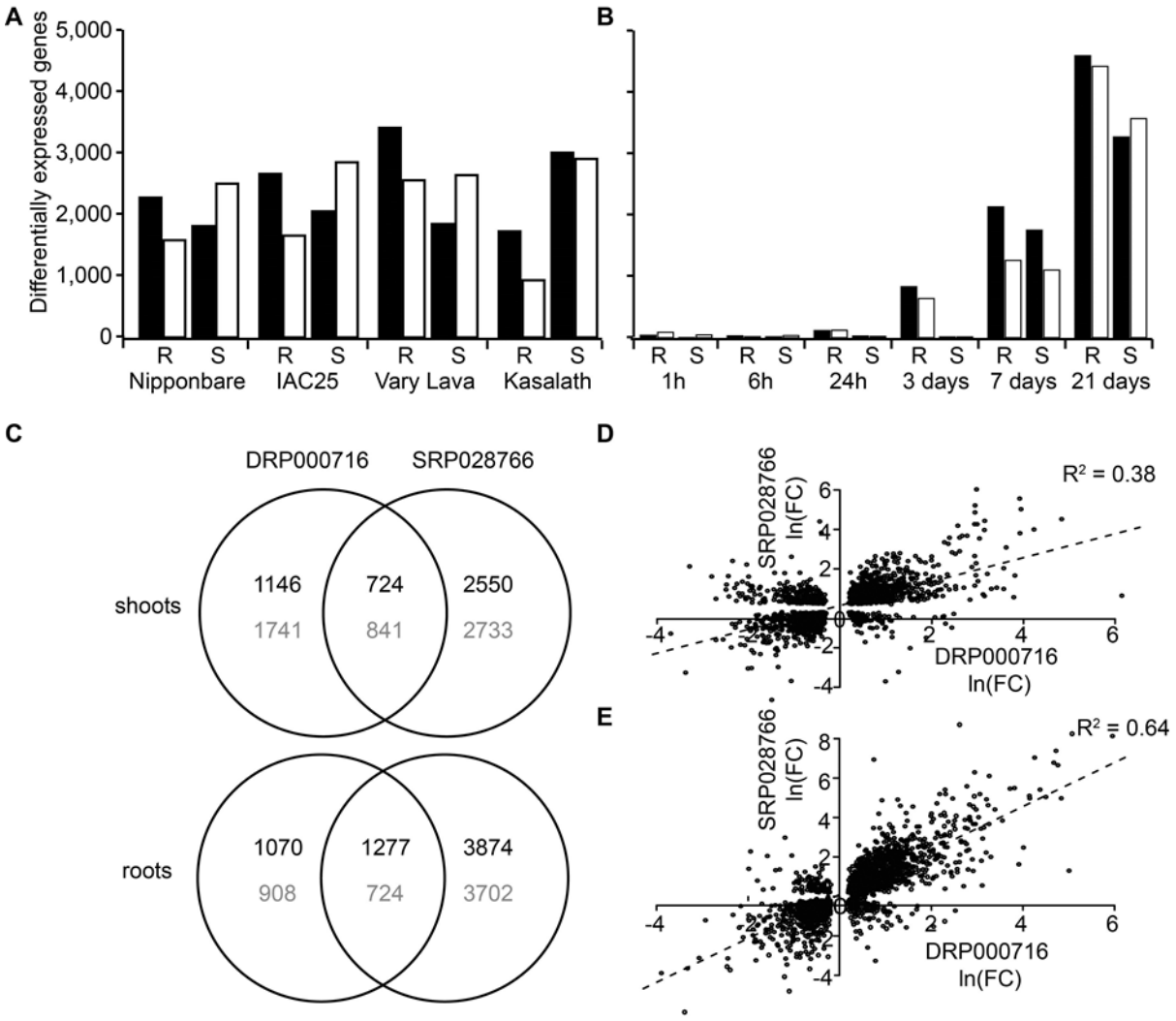

Figure S3. Comparison of differentially expressed genes identified in DRP00716 and SRP028766. (A) Genes differentially expressed after 22 days phosphate starvation in four rice cultivars. (B) Genes differentially expressed during a timecourse of phosphorous starvation in Nipponbare. In (A) and (B) filled bars represent upregulated genes, empty bars represent downregulated genes. (C) Differentially expressed genes identified in DRP000716 and SRP028766 at 22 and 21 days after phosphate starvation respectively in Nipponbare. Upregulated genes are shown in black, downregulated genes in grey. (D, E) Natural log (ln) fold change (FC) in genes differentially expressed under phosphate starvation in SRP028766 and DRP000716 in (D) shoots and (E) roots.

The integration of both studies with the same reference genes allows easy comparison between studies (above) and allows new hypotheses to be tested using existing data. We hypothesised that genes which are differentially expressed in roots of all four varieties from DRP000716 might be conserved phosphate responsive genes and should also be identified in SRP028766. To test this hypothesis we identified genes which were differentially expressed at 22 days in DRP000716 in all four varieties, and genes differentially expressed in each individual variety (Figure S4). For genes which were differentially expressed in individual varieties between 50 and 61 % were also differentially expressed in SRP028766 at 21 days (Figure S4). Amongst the genes conserved between all four varieties a higher percentage (81 %) were also detected in SRP028766: this suggests that not only are these genes conserved between varieties but they are also induced in independent experiments to a higher degree and would be strong candidates to investigate conserved phosphate responsive genes.

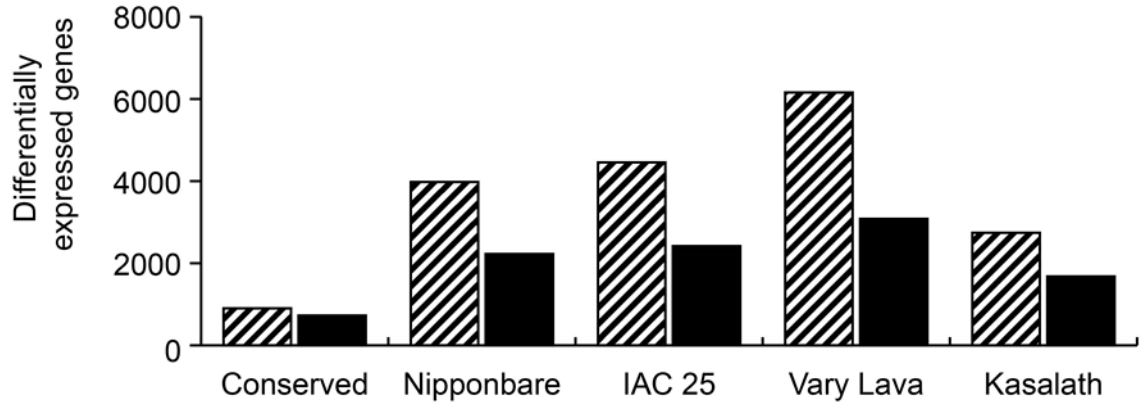

Figure S4. Intersection between genes differentially expressed in roots of all varieties and time-course expression. Differentially expressed genes identified in DRP000716 (striped bars) are also identified after 21 days phosphate starvation in SRP028766 (filled bars). Genes differentially expressed in all four varieties are called conserved.

Amongst these 726 conserved genes many have functions related to phosphate regulation including phosphate transporters (*PHT1-1*, *PHT1-4*, *PHT1-6*, *PHT1-8*, *PHT1-10*) and *SPX1*, *SPX2* and *SPX3* (Wang et al., 2009). Interestingly 91 genes have unknown functions and do not contain known interpro protein domains. An additional 32 genes contain domains of unknown function (DUF), one of which is represented in five genes: DUF581. In Arabidopsis

it has been proposed that DUF581 genes respond to specific environmental stresses and interact with SnRK1 to balance energy status (Nietzsche et al., 2014). These 5 genes (*Os03g0183500*, *Os04g0585700*, *Os06g0125200*, *Os06g0223700*, *Os09g0433800*) are upregulated 3-734 fold under phosphate starvation and may represent novel phosphate responsive genes.

Literature Cited

Liu C, Wu Y, Wang X (2012) bZIP transcription factor *OsbZIP52/RISBZ5*: a potential negative regulator of cold and drought stress response in rice. Planta **235:** 1157-1169

Nietzsche M, Schießl I, Börnke F (2014) The complex becomes more complex: protein-protein interactions of *SnRK1* with *DUF581* family proteins provide a framework for cell- and stimulus type-specific *SnRK1* signaling in plants. Frontiers in Plant Science **5**

Ohyanagi H, Tanaka T, Sakai H, Shigemoto Y, Yamaguchi K, Habara T, Fujii Y, Antonio BA, Nagamura Y, Imanishi T, Ikeo K, Itoh T, Gojobori T, Sasaki T (2006) The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. Nucleic Acids Research **34:** D741-D744

Oono Y, Kawahara Y, Yazawa T, Kanamori H, Kuramata M, Yamagata H, Hosokawa S, Minami H, Ishikawa S, Wu J, Antonio B, Handa H, Itoh T, Matsumoto T (2013) Diversity in the complexity of phosphate starvation transcriptomes among rice cultivars based on RNA-Seq profiles. Plant Molecular Biology **83:** 523-537

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, Orvis J, Haas B, Wortman J, Buell CR (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic Acids Research **35:** D883-D887

Secco D, Jabnoune M, Walker H, Shou H, Wu P, Poirier Y, Whelan J (2013) Spatio-Temporal Transcript Profiling of Rice Roots and Shoots in Response to Phosphate Starvation and Recovery. The Plant Cell **25:** 4285-4304

Wang Z, Hu H, Huang H, Duan K, Wu Z, Wu P (2009) Regulation of *OsSPX1* and *OsSPX3* on Expression of *OsSPX* domain Genes and Pi-starvation Signaling in Rice. Journal of Integrative Plant Biology **51:** 663-674

Zhang F, Sun Y, Pei W, Jain A, Sun R, Cao Y, Wu X, Jiang T, Zhang L, Fan X, Chen A, Shen Q, Xu G, Sun S (2015) Involvement of *OsPht1;4* in phosphate acquisition and mobilization facilitates embryo development in rice. The Plant Journal **82:** 556-569