Supplementary Data Set S14. Mapping strategy and assigning ORFs to species

Dupont et al. (Dupont et al., 2015) had made a Trinity (Haas et al., 2013) de novo assembly from all reads from all infected and uninfected tissue samples from our first biological replicate, yielding 300,938 transcripts, from which they extracted 126,580 predicted ORFs. Ninety percent of our reads mapped to these transcripts. Read distributions along transcripts and the frequent presence of regions matching several predicted ORFs or parts thereof on the same transcript (data not shown) suggested that some of the transcripts were fusions. In addition, 173,625 transcripts did not have BLAST hits against any of the predicted ORFs (13% of all reads mapped to these transcripts) and could be artifacts.

We found no evidence that any of our predicted ORFs were fusion artifacts, based on a BLAST search against *Oryza sativa* var. *japonica* gene models (Itoh et al., 2007), and *Epichloë festucae* gene models (Schardl et al., 2013), i.e. there were no instances in which a predicted ORF had significant similarity with several unrelated *O. sativa* or *E. festucae* models. We therefore based our transcriptome analyses on the mapping of reads (~65% of all reads) to the predicted ORFs.

Predicted ORFs with ≥95% sequence similarity to each other had been clustered by Dupont et al. (Dupont et al., 2015) and we used one representative ORF of each of the resulting 58,303 clusters for assigning ORFs to species. Of these 58,303 representative ORFs 8,105 had 95% identity over 100 bp and an E-value of < 0.05 in BLAST searches against published *E. festucae* gene models (Schardl et al., 2013). Few if any reads mapped to 8,096 of these representative ORFs in uninfected grass, and they were assigned to *E. festucae*. Also assigned to *E. festucae* were 40 additional ORFs to which (i) high numbers of reads mapped in several infected tissue but not in uninfected plants and for which (ii) BLAST searches suggested that they were not grass-derived. Thus 8,136 ORFs were assigned to *E. festucae*. Subsequently FASTA v.36.3.6 (Pearson et al., 1997) was used to identify likely *E. festucae* gene model (Schardl et al., 2013) homologs on the basis of ≥95% identity over 100 bp and an E-value of < 10^{-5}; the procedure was applied not only to these 8,136 representative ORFs but to all ORFs of the clusters they represented.

All 50,198 representative ORFs not assigned to *E. festucae* had been subjected to BLAST searches against nonredundant protein and nucleotide NCBI databases, and 45,514 had produced significant hits (Dupont et al., 2015). Of these, 9618 ORFs were assigned to *L. perenne* because BLAST hits were restricted to five *L. perenne*-related species to which ORFs mapped most often (*Oryza sativa* subsp*. japonica*, *Brachypodium distachyon*, *Sorghum bicolor*, *Zea mays*, *Oryza sativa* subsp*. indica*, *Hordeum sativum*). A further 35,892 had hits to other species as well. Of these, 32,397 were assigned to *L. perenne*, because (i) BLAST hits indicated an *L. perenne* origin and/or (ii) reads mapped to the majority of uninfected *L. perenne* tissues and/or (iii) a high number of reads mapped to one or several uninfected above-ground *L. perenne* tissues. (Of these 32,397 ORFs, 26,974 had reads mapping to them in 21 or more of all 42 tissues collected in the sampling of three biological

replicates; the remaining 5423 ORFs had multiple BLAST hits to *Poaceae* and/or spermatophytes.)

Out of the 4,684 representative ORFs that produced no significant hits, 68 were also assigned to *L. perenne*. They are most likely part of the grass genome rather than those of other unknown species present in some tissues, because reads mapped to them in a multitude of uninfected tissues.

Based on the species assignments of the 50,198 representatives, out of all 126,580 ORFs 96,708 were assigned to *L. perenne* and 9,567 to *E. festucae* (Species assignment of individual ORFs is shown in Supplementary Data Set S16). We could not confidently assign 20,305 ORFs to either species (Supplementary Data Set S16). BLAST results suggest that many of these are derived from other species present, such as members of the root microbiome, epiphytic fungi, arthropods and viruses (Supplementary Data Set S17).

In infected grass, of all reads that mapped to ORFs, 98.7% were captured by ORFs assigned to *L. perenne*, 0.8% by ORFs assigned to *E. festucae*, and 0.4% mapped to ORFs not assigned to either species. In several tissues the number of reads mapping to unassigned ORFs was ≥2 fold higher than the number of reads mapping to *Epichloë*-assigned ORFs (Table S1). Incorrect assignment of ORFs to *Epichloë* could thus have seriously interfered with analyses of the *Epichloë* transcriptome. However, the chance that ORFs were wrongly assigned to *Epichloë* was low, since 99.5% of *Epichloë*-assigned ORFs matched published gene models (see above). Also, the ratio between reads mapping to unassigned ORFs and reads mapping to *L. perenne*-assigned ORFs was slightly lower in infected than in uninfected tissues (averages $4.4 \times 10^{-3} \pm 1.5 \times 10^{-3}$ versus $4.8 \times 10^{-3} \pm 1.2 \times 10^{-3}$; it was lower in uninfected plants in 15/21 comparisons of the same tissue types from the same biological replicate; Table S1); the ratio should have been higher in infected tissue if we had failed to assign to *Epichloë* a significant number of *Epichloë* ORFs from our de novo assembly.

Literature Cited

Dupont, P.-Y., Eaton, C.J., Wargent, J.J., Fechtner, S., Solomon, P., Schmid, J., Day, R.C., Scott, B., and Cox, M.P. 2015. Fungal endophyte infection of ryegrass reprograms host metabolism and alters development. New Phytol. 208:1227–1240.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., Macmanes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N., and Regev, A. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8:1494-1512.

Itoh, T., Tanaka, T., Barrero, R.A., Yamasaki, C., Fujii, Y., Hilton, P.B., Antonio, B.A., Aono, H., Apweiler, R., Bruskiewich, R., Bureau, T., Burr, F., Costa de Oliveira, A., Fuks, G., Habara, T., Haberer, G., Han, B., Harada, E., Hiraki, A.T., Hirochika, H., Hoen, D., Hokari, H., Hosokawa, S., Hsing, Y., Ikawa, H., Ikeo, K., Imanishi, T., Ito, Y., Jaiswal, P., Kanno, M., Kawahara, Y., Kawamura, T., Kawashima, H., Khurana, J.P., Kikuchi, S., Komatsu, S., Koyanagi, K.O., Kubooka, H., Lieberherr, D., Lin, Y.-C., Lonsdale, D., Matsumoto, T., Matsuya, A., McCombie, W.R., Messing, J., Miyao, A., Mulder, N., Nagamura, Y., Nam, J., Namiki, N., Numa, H., Nurimoto, S., O'Donovan, C., Ohyanagi, H., Okido, T., Oota, S., Osato, N., Palmer, L.E., Quetier, F., Raghuvanshi, S., Saichi, N., Sakai, H., Sakai, Y., Sakata, K., Sakurai, T., Sato, F., Sato, Y., Schoof, H., Seki, M., Shibata,

M., Shimizu, Y., Shinozaki, K., Shinso, Y., Singh, N.K., Smith-White, B., Takeda, J.-i., Tanino, M., Tatusova, T., Thongjuea, S., Todokoro, F., Tsugane, M., Tyagi, A.K., Vanavichit, A., Wang, A., Wing, R.A., Yamaguchi, K., Yamamoto, M., Yamamoto, N., Yu, Y., Zhang, H., Zhao, Q., Higo, K., Burr, B., Gojobori, T., Sasaki, T., and for the Rice Annotation, P. 2007. Curated genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana*. Genome Res. 17:175-183.

Pearson, W.R., Wood, T., Zhang, Z., and Miller, W. 1997. Comparison of DNA sequences with protein sequences. Genomics 46:24-36.

Schardl, C.L., Young, C.A., Hesse, U., Amyotte, S.G., Andreeva, K., Calie, P.J., Fleetwood, D.J., Haws, D.C., Moore, N., Oeser, B., Panaccione, D.G., Schweri, K.K., Voisey, C.R., Farman, M.L., Jaromczyk, J.W., Roe, B.A., O'Sullivan, D.M., Scott, B., Tudzynski, P., An, Z., Arnaoudova, E.G., Bullock, C.T., Charlton, N.D., Chen, L., Cox, M., Dinkins, R.D., Florea, S., Glenn, A.E., Gordon, A., Güldener, U., Harris, D.R., Hollin, W., Jaromczyk, J., Johnson, R.D., Khan, A.K., Leistner, E., Leuchtmann, A., Li, C., Liu, J., Liu, J., Liu, M., Mace, W., Machado, C., Nagabhyru, P., Pan, J., Schmid, J., Sugawara, K., Steiner, U., Takach, J.E., Tanaka, E., Webb, J.S., Wilson, E.V., Wiseman, J.L., Yoshida, R., and Zeng, Z. 2013. Plant-Symbiotic Fungi as Chemical Engineers: Multi-Genome Analysis of the Clavicipitaceae Reveals Dynamics of Alkaloid Loci. PLoS Genet 9:e1003323.