Substrate-Specific Differential Gene Expression and RNA editing in the Brown Rot Fungus *Fomitopsis pinicola*

Baojun Wu[a], Jill Gaskell[b], Benjamin W. Held[c], Cristina Toapanta[c], Thu Vuong[d], Steven Ahrendt[e,f], Anna Lipzen[e], Jiwei Zhang[g], Jonathan S. Schilling[g], Emma Master[d], Igor V. Grigoriev[e,f], Robert A. Blanchette[c], Dan Cullen[b], David S. Hibbett[a]#

[a]Biology Department, Clark University, Worcester MA, USA
[b]USDA Forest Products Laboratory, Madison, WI, USA
[c]Department of Plant Pathology, University of Minnesota, St. Paul MN, USA
[d]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto ON, Candada
[e]Department of Energy Joint Genome Institute, Walnut Creek CA, USA
[f]Department of Plant and Microbial Biology, University of California, Berkeley CA, USA
[g]Department of Plant and Microbial Biology, University of Minnesota, St. Paul MN, USA

SUPPLEMENTAL MATERIAL

Time-of-Flight Secondary Ion Mass Spectometry (ToF-SIMS)

Methods

Positive ion ToF-SIMS spectra were obtained using a ToF-SIMS V instrument with a bismuth liquid metal ion source and a reflectron type mass analyzer (IONTOF GmbH, Münster, Germany). At least eight spectra were acquired for each sample, using 50 keV Bi_3^{++} primary ions with a pulsed current of approximately 0.3 pA. Each spectrum covers a 500 x 500 µm area using 128 x 128 pixels in a random raster pattern. The cycle time is 100 µs and ion doses were kept below 1 x 10^{12} ions/cm^2 to limit sample damage. Low-energy electron flooding (20 eV) was used to charge compensation. Ion spectra were calibrated to CH_3^+, $C_2H_3^+$ and $C_3H_5^+$ ions using SurfaceLab v.6 software.

Principal component analysis (PCA) was performed using Matlab R2014a (The Mathworks Inc., Massachusetts, USA) with PLS Toolbox v.7 (Eigenvector Research Inc., Washington, USA). ToF-SIMS spectra were preprocessed by first normalizing peak intensities to the total ion intensity to eliminate the influence of overall intensity changes that might arise from topography or variations in instrumental setup, and then by mean centering so that all principal components described variations from the data set mean.

Results

To directly investigate fungal action on pine and spruce, residual wood wafers after *F. pinicola* treatments were recovered at day 10 and day 30, and analyzed by ToF-SIMS. Peaks corresponding to protein, extractives, lignin and polysaccharides (1, 2) were extracted from spectra, and their intensities between wood treatments were compared using principal component analysis (PCA). The fungal treated wood samples, both day 10 and day 30, were separated from the corresponding controls by protein-specific peaks (Fig. S3), revealing the presence of proteins and enzymes on the surface of treated wood samples. Another major difference between the treated samples and the controls was the depletion of extractive peaks in the treated samples (Fig. S3). A closer look at PCA loadings also indicated the removal of generic aromatic peaks at 65 Da and 77 Da (Fig. S3). These modifications are in agreement with one of the functional roles of P450s in the degradation of complex wood extractives (3-5).

Some levels of polysaccharide degradation, particularly in pine, were observed: The extractive peak of 43 Da, which was also ascribed to acetyl groups, probably from hemicellulose degradation (6), was depleted much more in treated pine. Two polysaccharide peaks, at 99 and 113 Da, were slightly depleted in treated pine, supporting polysaccharide modifications in pine.

Supplemental figures

Figure S1. Hierarchical clustering of RNA-seq samples from three different substrates in submerged culture at day 5. Gene expression profiles were clustered using normalized counts of gene expression (FPKM).

Figure S2. Comparison of RNAseq and qRTPCR for two differentially regulated genes. Primers for peptidase (Protein ID Fompi_1113467) and thaumatin (Fompi_110401) amplification flanked introns yielding cDNA/gDNA amplicons of 461/627nt and 214/320nt, respectively. Poly(A) RNA was reverse transcribed with GoScript (Promega, Madison, WI) and qRT-PCR performed as described [(7-11) reviewed in (12)]. Poly(A) RNA was isolated by magnetic capture using oligo $(dT)_{25}$ Dynabeads (Oslo, Norway) as recommended by the manufacturer, and equivalency between the genomic and cDNA amplicons were estimated using Syngene software (Cambridge, UK). Peptidase primers were 5'-GGCGAAGAGCGACGGGAACAAACT and 5'-CCATCACCCGAGGCGAATAGAACA, and thaumatin primers were 5' CACGGCCCCGCTGGACACCTA and 5'-CGCCGCAGCCATTACCCTCGTTA.

Figure S3. PCA analysis of ToF-SIMS spectra obtained from residual spruce wafers (A, B) and pine wafers (C, D) after *F. pinicola* incubation. PCA scores (A, C) and loadings (B, D) show that the separation of treatment groups was explained by differences in protein and extractives ions. Score plots: D10 - day 10; D30 - day 30; blue - treated samples, red - control samples. Loadings plots: green circles - protein peaks; white diamonds - extractives peaks; red squares - polysaccharide peaks, and blue stars - lignin peaks. At least eight spectra were acquired for each sample.

Figure S4. Distribution of RNA-editing loci and RNA-editing types at two time points (day 10 and day 30) on two substrates (pine and spruce).

Figure S5. Hierarchical clustering of RNA-seq samples from different substrates, time points and culture methods. The 5-day samples were from submerged culture, while the 10-day and 30-day samples were from wood wafers. Gene expression profiles were clustered using the normalized counts of gene expression (FPKM).

Supplemental tables

Table S1. Transcriptome expression levels of all *Fomitopsis pinicola* genes after 5-day. submerged culture**.**

Table S2. Position and type of all RNA editing events.

Table S3. Transcriptome expression levels of all *Fomitopsis pinicola* genes after 10-day and 30-day wood wafer culture.

Table S4. Distribution in relative proportion of wood-decay CAZymes, Redox enzyme and Cytochrome_P450.

Table S5. Summary of mapping reads from each sample.

REFERENCES

1. Goacher RE, Edwards EA, Yakunin AF, Mims CA, Master ER. 2012. Application of time-of-flight-secondary ion mass spectrometry for the detection of enzyme activity on solid wood substrates. Anal Chem 84:4443-4451.
2. Goacher RE, Tsai AY, Master ER. 2013. Towards practical time-of-flight secondary ion mass spectrometry lignocellulolytic enzyme assays. Biotechnol Biofuels 6:132.
3. Martinez D, Challacombe J, Morgenstern I, Hibbett D, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Duenas FJ, Martinez AT, Kersten P, Hammel KE, Vanden Wymelenberg A, Gaskell J, Lindquist E, Sabat G, Bondurant SS, Larrondo LF, Canessa P, Vicuna R, Yadav J, Doddapaneni H, Subramanian V, Pisabarro AG, Lavin JL, Oguiza JA, Master E, Henrissat B, Coutinho PM, Harris P, Magnuson JK, Baker SE, Bruno K, Kenealy W, Hoegger PJ, Kues U, Ramaiya P, Lucas S, Salamov A, Shapiro H, Tu H, Chee CL, Misra M, Xie G, Teter S, Yaver D, James T, Mokrejs M, Pospisek M, Grigoriev IV, Brettin T, et al. 2009. Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci U S A 106:1954-1959.
4. Doddapaneni H, Chakraborty R, Yadav JS. 2005. Genome-wide structural and evolutionary analysis of the P450 monooxygenase genes (P450ome) in the white rot fungus *Phanerochaete chrysosporium*: evidence for gene duplications and extensive gene clustering. BMC Genomics 6:92.
5. Doddapaneni H, Yadav JS. 2005. Microarray-based global differential expression profiling of P450 monooxygenases and regulatory proteins for signal transduction pathways in the white rot fungus *Phanerochaete chrysosporium*. Mol Genet Genomics 274:454-466.
6. Goacher RE, Michienzi CL. 2018. How does soaking wood in various pH buffers impact ToF-SIMS spectra. Surface and Interface Analysis 50:163-170.
7. Lamar RT, Schoenike B, Vanden Wymelenberg A, Stewart P, Dietrich DM, Cullen D. 1995. Quantitation of fungal mRNAs in complex substrates by reverse transcription PCR and its application to *Phanerochaete chrysosporium*-colonized soil. Appl Environ Microbiol 61:2122-2126.
8. Bogan B, Schoenike B, Lamar R, Cullen D. 1996. Expression of *lip* genes during growth in soil and oxidation of anthracene by *Phanerochaete chrysosporium.* Appl Environ Microbiol 62:3697-3703.
9. Janse BJH, Gaskell J, Akhtar M, Cullen D. 1998. Expression of *Phanerochaete chrysosporium* genes encoding lignin peroxidases, manganese peroxidases, and glyoxal oxidase in wood. Appl Environ Microbiol 64:3536-3538.
10. Vanden Wymelenberg A, Denman S, Dietrich D, Bassett J, Yu X, Atalla R, Predki P, Rudsander U, Teeri TT, Cullen D. 2002. Transcript analysis of genes encoding a family 61 endoglucanase and a putative membrane-anchored family 9 glycosyl hydrolase from *Phanerochaete chrysosporium*. Appl Environ Microbiol 68:5765-5768.
11. Vanden Wymelenberg A, Sabat G, Mozuch MD, Kersten P, Cullen D, Blanchette RA. 2006. Structure, organization, and transcriptional regulation of a family of copper radical oxidase genes in the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Appl Environ Microbiol 72:4871-4877.

12. Zentilin L, Giacca M. 2007. Competitive PCR for precise nucleic acid quantification. Nat Protoc 2:2092-2104.

Fig. S1

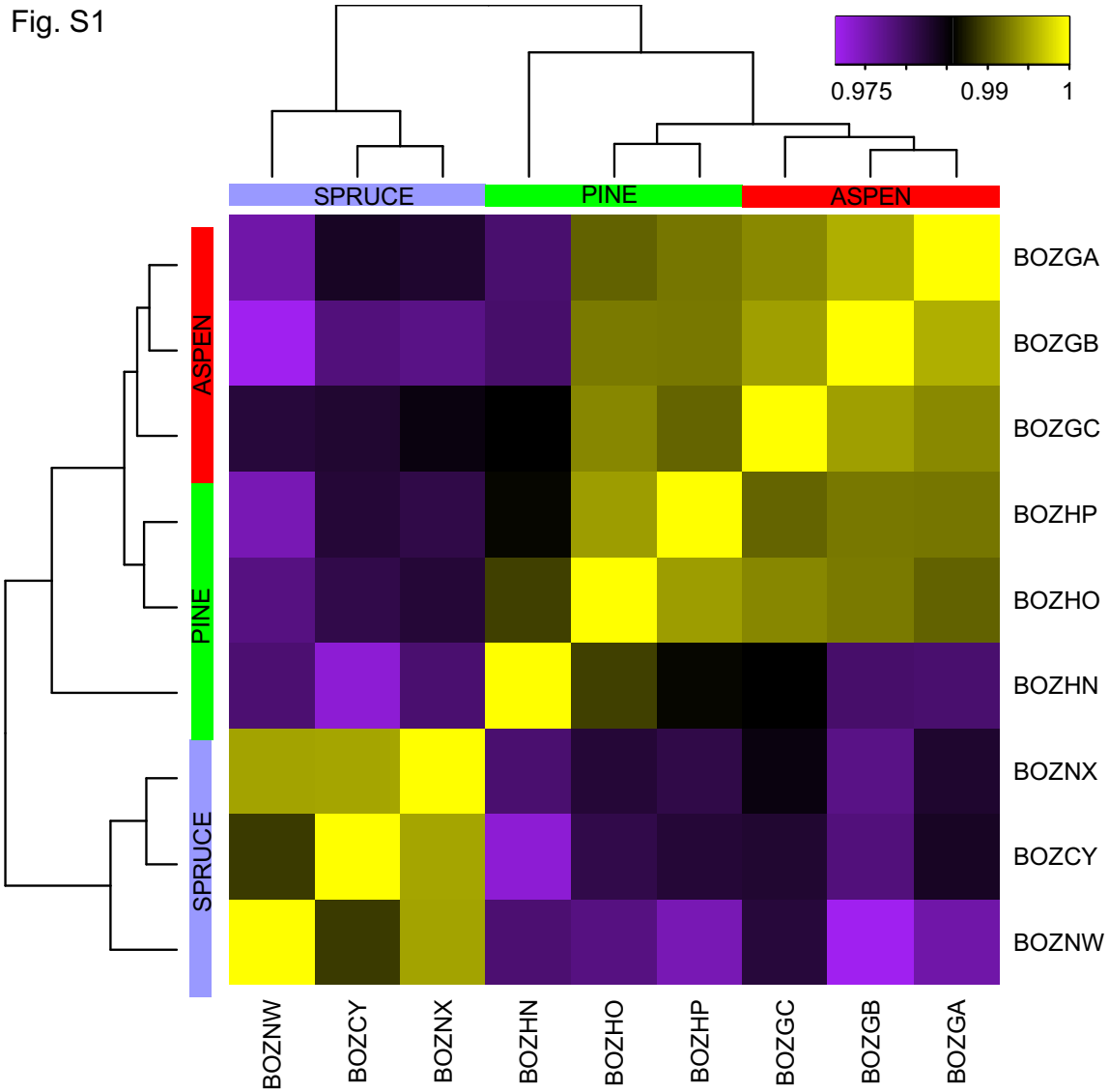

0.975
0.99
1
SPRUCE
PINE
ASPEN
ASPEN
PINE
SPRUCE
BOZGA
BOZGB
BOZGC
BOZHP
BOZHO
BOZHN
BOZNX
BOZCY
BOZNW
BOZNW
BOZCY
BOZNX
BOZHN
BOZHO
BOZHP
BOZGC
BOZGB
BOZGA

Fig. S2

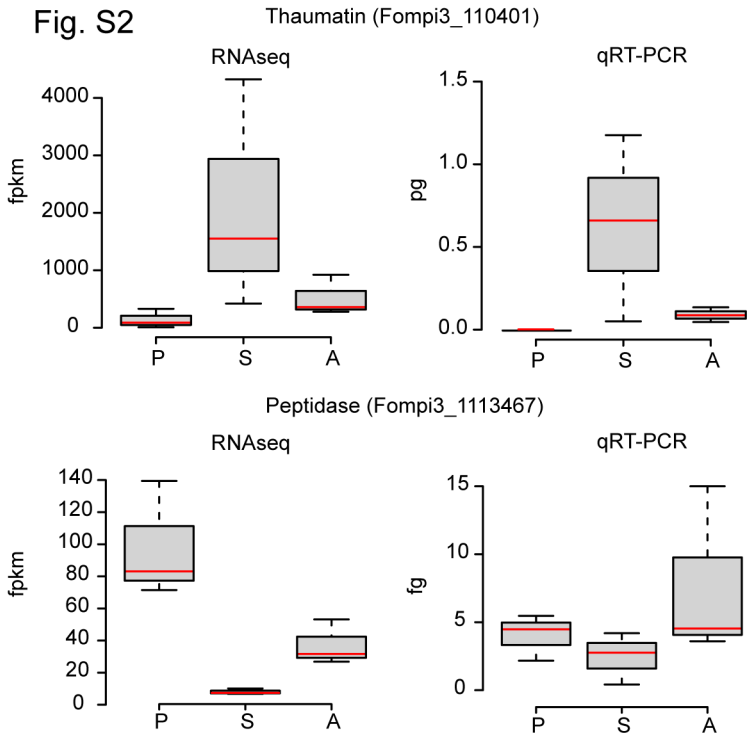

Thaumatin (Fompi3_110401)
RNAseq
qRT-PCR
fpkm
pg
4000
3000
2000
1000
0
1.5
1.0
0.5
0.0
P
S
A
Peptidase (Fompi3_1113467)
RNAseq
qRT-PCR
fpkm
fg
140
120
100
80
60
40
20
0
15
10
5
0
P
S
A

Fig. S3

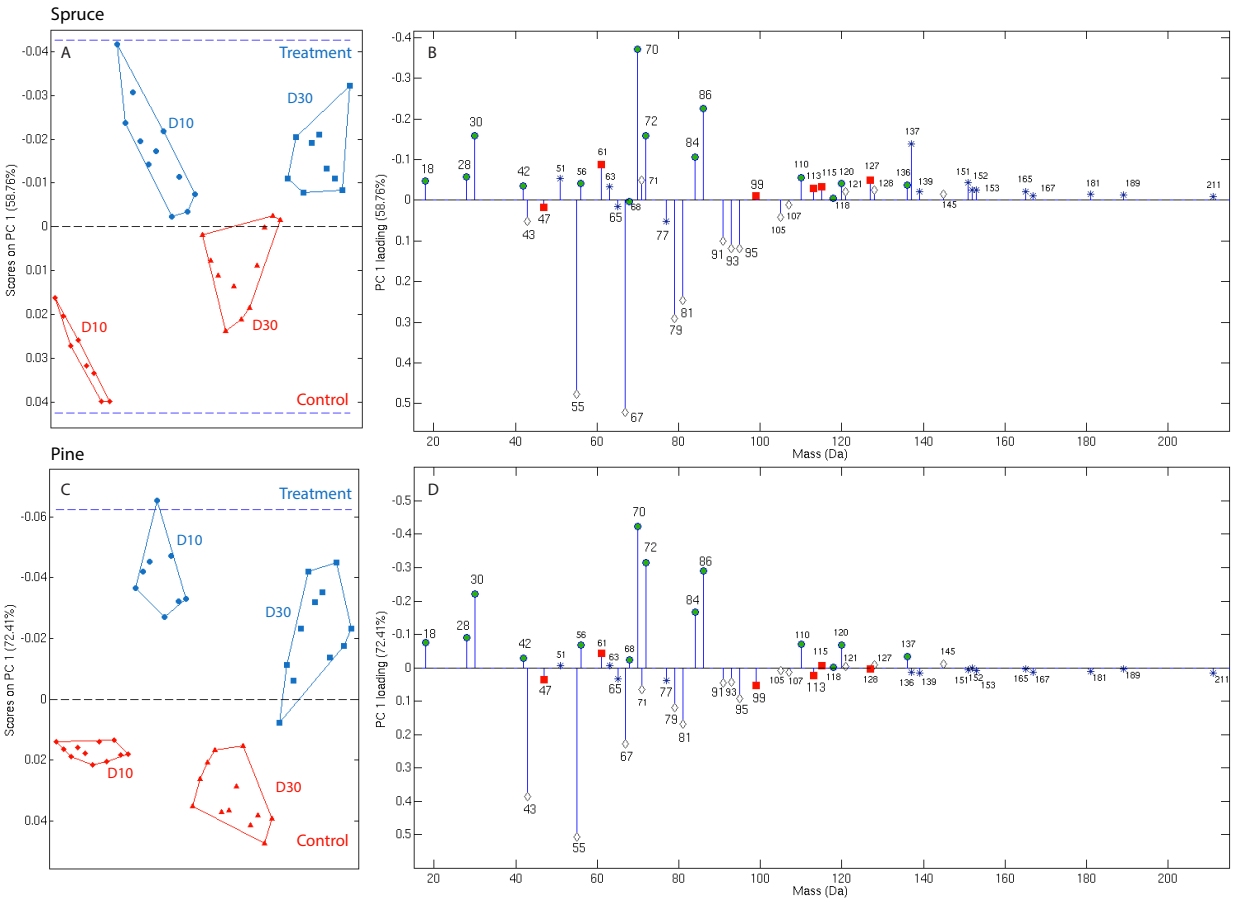

Spruce
A
Treatment
D10
D30
D10
D30
Control
Scores on PC 1 (58.76%)
B
PC 1 loading (58.76%)
Mass (Da)
Pine
C
Treatment
D10
D30
D10
D30
Control
Scores on PC 1 (72.41%)
D
PC 1 loading (72.41%)
Mass (Da)

Fig. S4

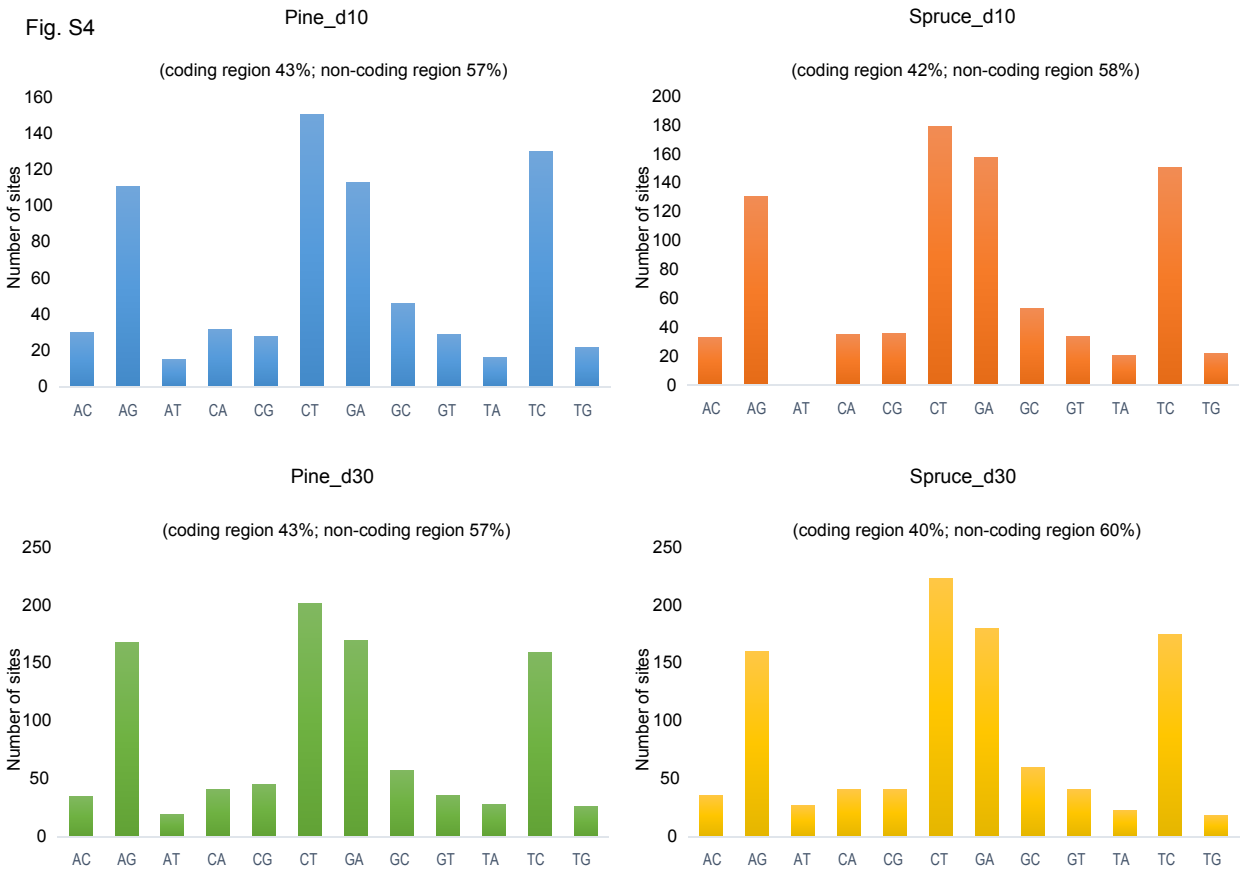
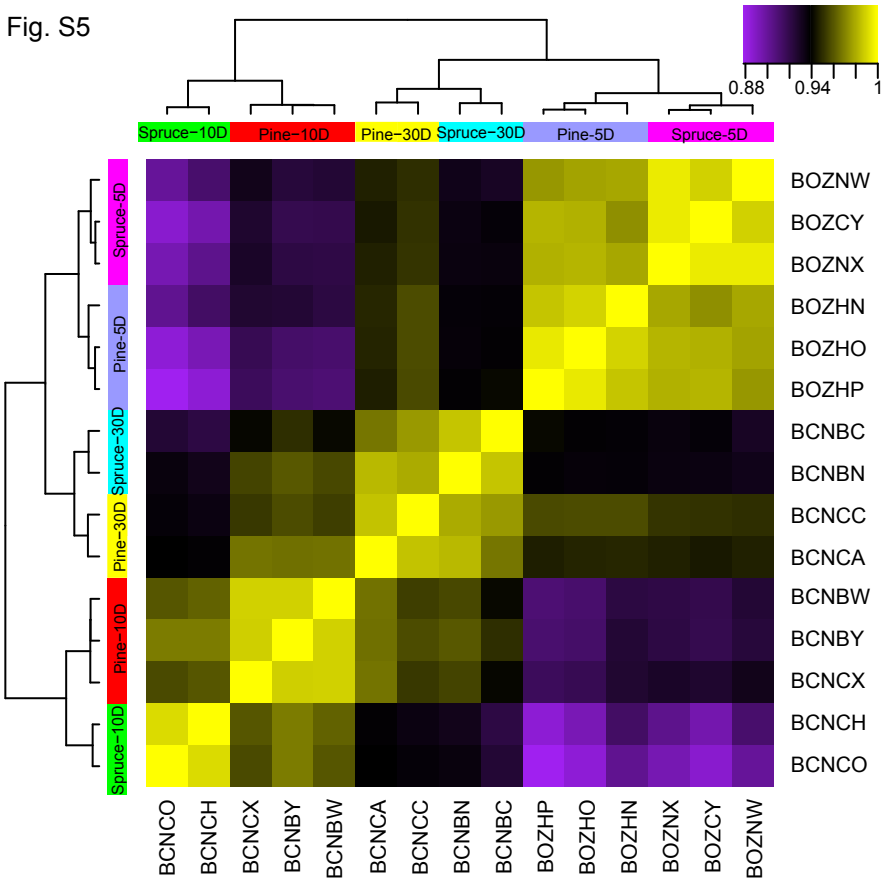

Pine_d10
(coding region 43%; non-coding region 57%)
Number of sites
0
20
40
60
80
100
120
140
160
AC AG AT CA CG CT GA GC GT TA TC TG
Spruce_d10
(coding region 42%; non-coding region 58%)
Number of sites
0
20
40
60
80
100
120
140
160
180
200
AC AG AT CA CG CT GA GC GT TA TC TG
Pine_d30
(coding region 43%; non-coding region 57%)
Number of sites
0
50
100
150
200
250
AC AG AT CA CG CT GA GC GT TA TC TG
Spruce_d30
(coding region 40%; non-coding region 60%)
Number of sites
0
50
100
150
200
250
AC AG AT CA CG CT GA GC GT TA TC TG

Fig. S5


0.88
0.94
1
Spruce−10D
Pine−10D
Pine−30D
Spruce−30D
Pine−5D
Spruce−5D
BOZNW
BOZCY
BOZNX
BOZHN
BOZHO
BOZHP
BCNBC
BCNBN
BCNCC
BCNCA
BCNBW
BCNBY
BCNCX
BCNCH
BCNCO