Supplemental Methods

Beetle/fungal collections and DNA extractions.

To generate pure cultures for the fungi, we isolated *Ceratocystiopsis* from single spores and *Entomocorticium* from hyphal tips and the resulting mycelia were initially cultured on 2% MEA and then transferred to liquid media (2% malt extract and sterile water) to generate enough material for DNA extraction and sequencing. Bark beetles used as outgroups (*D. approximatus*, *D. frontalis*, and *D. adjunctus*) were also collected in the summer of 2011. *Dendroctonus approximatus* and *D. adjunctus* were collected near Beaver, UT (38° 22' N, 112° 31' W) while *D. frontalis* was collected near Morenci, AZ (33° 28' N, 109° 22' W). Female beetle tissue used in DNA extractions included the head, the posterior portion of the pronotum (therefore excluding the mycangia) and the anterior portion of the thorax (excluding elytra and gut). Male beetles lack a mycangium and therefore DNA was extracted from the head to the anterior portion of the thorax.

Genome sequencing and assembly

To generate ALLPATHS-LG [1] assemblies for each species, we generated overlapping read libraries (Illumina 100bp PE library with insert size of ~180 bp), and a mate pair libraries (Illumina 100bp PE with insert size of ~3-4 kb). Sequencing was performed by GENEWIZ (South Plainfield, NJ), the University of Utah Microarray and Genomic Analysis Core Facility (Salt Lake City, UT), University of Montana Genomics Core Facility (Missoula, MT) and Vincent J. Coates Genomics Sequencing Laboratory, Berkeley, CA.

For each species, all raw reads (both mate pair and overlapping) were first adapter trimmed and quality trimmed using default settings in SeqyClean version 1.8.10. An initial assembly was performed using overlapping reads with MaSuRCA version 3.1.0 [2]. We then

used the Nextclip pipeline version 0.7 [3], to process and map mate-pair data and estimate insert size. Properly oriented mate pair reads were then used along with cleaned short reads to produce a final assembly with ALLPATHS-LG (release R48777) with basic evaluation implemented. For the western pine beetle and basidiomycete genome, setting the parameter HAPLOIDIFY=True greatly improved the assembly.

RADseq library preparation

Approximately 200 ng total DNA for each beetle and ~300ng for each fungal isolate was used for RADseq. Each species was prepared separately in their own set of libraries (8-12 individuals per library). We followed our own single restriction enzyme digest protocol which is similar to other single restriction enzyme digest protocols (e.g., Etter et al. [4]), yet uses AMPure XP bead (Beckman Coulter, Brea CA) cleanups. We also used a slightly modified PE2 adapter (5'-ATCGGAAGAGCGAGAACAA-3', 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT- 3') which allows for combinatorial inline and third read barcoding of samples [5]. In short, DNA was cleaned, digested with PstI (New England Biolabs), adapters with barcodes were ligated to the restriction cut sites, and samples were then pooled and sonicated using a Bioruptor (Diagenode Inc., Sparta, NJ). After shearing, DNA was blunt-end repaired, A-tailed and a small y-adapter (above) added. The indexing read barcode and Illumina specific adapter sequences were then added via PCR which was done in 5 separate 20 ul reactions for each library using the following conditions: 98°C for 30s: 12-14X (98°C for 10s, 65°C for 30s, 72°C for 30s) with a final extension at 72°C for 5 minutes. An appropriate cycle number (12-14) for each library was determined via qPCR. Libraries from each species were pooled and sequenced on an Illumina HiSeq 2000 with 100 bp

paired end reads at either the University of Utah Microarray and Genomic Analysis Core Facility (Salt Lake City, UT) or the Vincent J. Coates Genomics Sequencing Laboratory (Berkeley, CA).

RADseq read mapping and SNV identification

All raw Illumina reads were processed using the process_radtags program from STACKS [6]. We mapped reads using BWA-MEM, [7] and manipulated reads using samtools [8] and Picardtools. We removed PCR duplicates using samtools rmdup and generated summary statistics using samtools flagstat and the GATK's DepthOfCoverage tool. We genotyped all *Ceratocystiopsis*, *Entomocorticium*, and *Dendroctonus* samples in genus specific runs using GATK UnifiedGenotyper version 3.1-1 and output mode EMIT_ALL_SITES. Because of the larger number of *Dendroctonus* samples and computational limitations, we randomly assigned ~40 beetles to each genotyping run, and then merged the VCFs using the GATK's CombineVariants tool. SNV filtering was performed using VCFtools version 0.1.12b [9].

Beetle data filtering

We first used BLAST to compare all western pine beetle scaffolds to the mountain pine beetle genome [10] and retained beetle specific scaffolds (top BLAST hit > 1kb). We flagged all scaffolds that hit any putative mountain pine beetle X-linked scaffold and also compared male and female RADseq coverage over western pine beetle scaffolds to identify additional putative X-linked scaffolds. Any scaffold where normalized female coverage was >1.5× male coverage was also flagged (Supplemental Fig. 3) and all flagged putative X-linked regions were excluded from downstream analyses.

Due to lower and more variable coverage of our beetle samples (see RESULTS), we filtered to create datasets suited to particular analyses. For western pine beetle population genetic structure analyses, we removed all outgroups from our initial VCF (above), excluded sites that did not pass GATK's quality filter and filtered out genotypes with quality < 30. We further excluded any position that deviated significantly from Hardy-Weinberg (HW) proportions in any one of the seven populations where we had data for seven or more individuals. We then filtered to only include sites where 70% of individuals had a genotype call, and genotype calls with a minimum depth < 5 and maximum depth > 200. We then only included biallelic sites and sites with a minor allele frequency (MAF) > 0.10. To estimate divergence, diversity (π) and Tajima's D, we used the 28 highest coverage East and West individuals and filtered our original VCF (above) using only the genotype quality and coverage filters (above) to maximize our confidence in genotype calls at both variant and invariant positions. For our phylogenetic dataset we selected the 55 highest coverage western pine beetles (SNV average coverage > 15×) and the outgroup *D. approximatus* and removed sites that failed the GATK's filters and excluded genotype calls with quality < 30 and depth > 200.

Basidiomycete data filtering

To mask putative problematic regions of the highly repetitive *E.* sp. B genome (Supplemental Table 1) we re-mapped reads used for assembly and estimated median coverage to be ~68×. Regions with reduced (< 30×) or excessive coverage (> 110×) were masked. Further, lineage specific (*E.* sp. B_1, *E.* sp. B_2, and *E.* sp. B_3) paralogues were also identified from our RADseq data (below) and masked by identifying regions where > 50% of all calls were erroneous heterozygous calls. This removed an additional 73,616 bp of sequence (8% of the

total). After these two genomic filters, only 3,656 heterozygous positions remained (0.01% of all genotype calls), which was comparable to analyses of the other fungal RADseq data (below) and seemed sufficiently low.

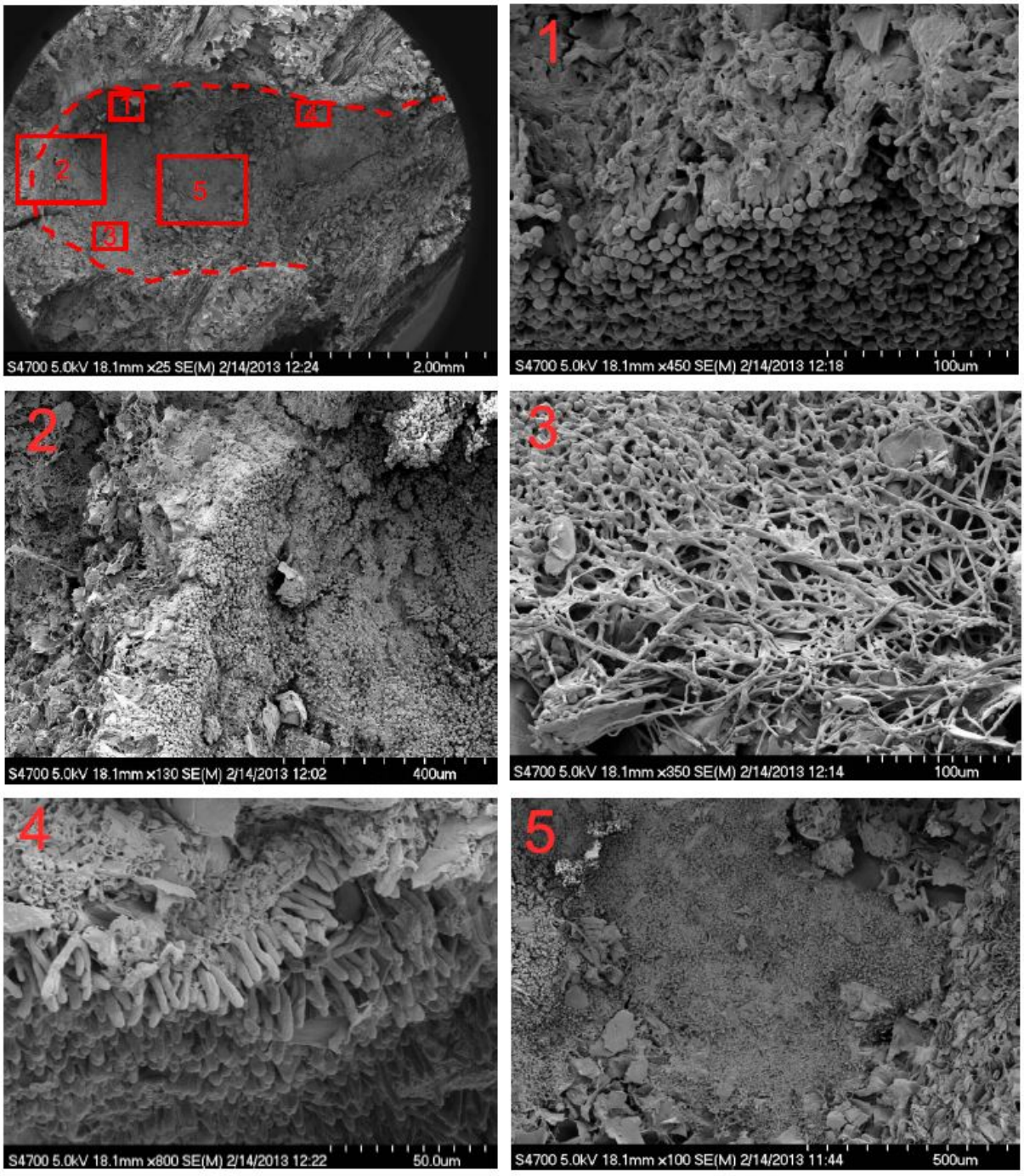
For filtering we excluded sites that did not pass GATK's quality filters and/or where depth was < 10× for any one individual. SNV's with genotype quality < 30 were changed to unknown, and we allowed for triallelic positions. After these filters, only a small percentage (0.01%) of all genotyping calls were still incorrect heterozygous calls, and those were all changed to missing data. We concatenated all genotype calls for all individuals resulting in 315,655 bp of concatenated sequence. For population genetic analyses we refiltered the data by first removing outgroups and then following the steps outlined above but only allowed biallelic SNVs. This resulted in 1,328,866 bp of total sequence.

Ascomycete data filtering

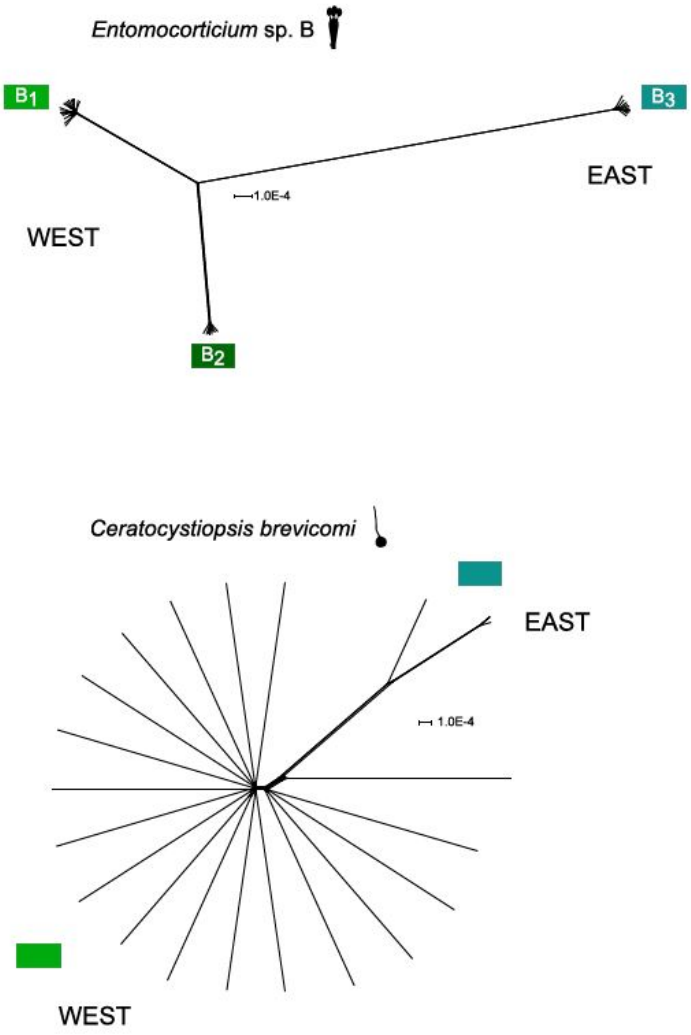
We first removed the outgroup (*C. ranaculosus*) from the VCF and then excluded sites that did not pass the GATK quality filters and/or where depth was < 20× for any one individual. Then, SNV calls with a genotype quality < 30 were changed to unknown and only sites with no more than 2 alleles were kept. After these filters, only a small fraction (0.02%) of all genotyping calls were still incorrect heterozygous calls and those were all changed to missing data. In total, this resulted in genotype calls for 2,926,074 bases in every *C. brevicomi* individual. We then returned to the outgroup (*C. ranaculosus*) and filtered to only include positions genotyped (variant or invariant) in *C. brevicomi,* using the same depth and genotype quality filters described above but allowing for triallelic sites. VCFs of the variant + invariant, or just SNVs were then used for downstream analyses.

Scanning electron microscope imaging

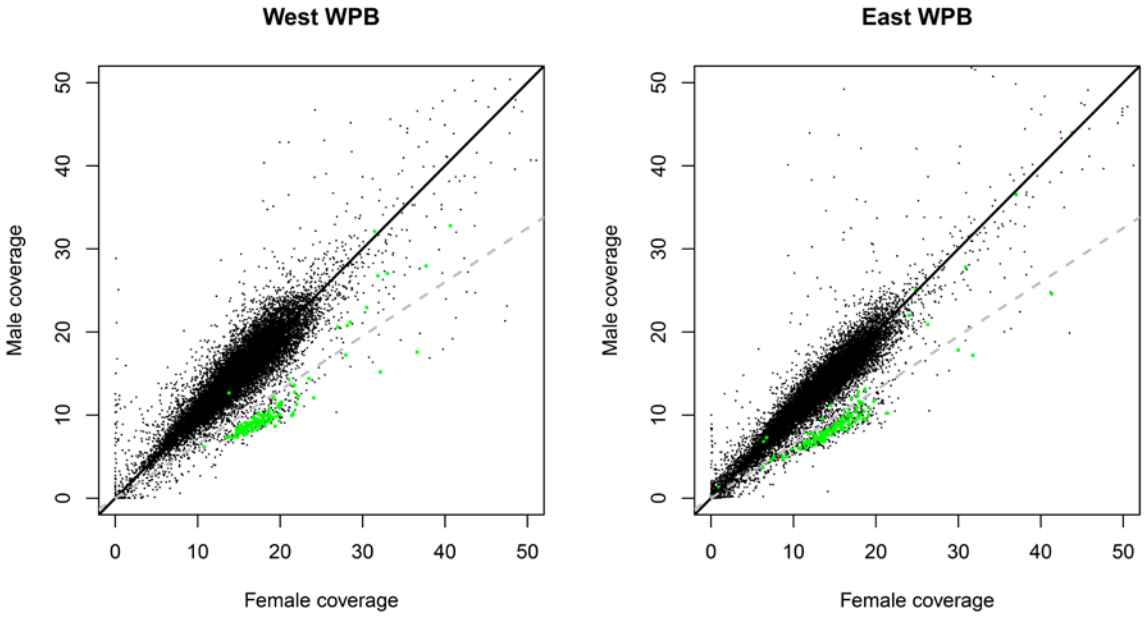
A ponderosa pine infested with western pine beetle was collected on October of 2012 near the LF site (Supplemental Table 3). The outer bark was fractured to expose larvae and larval tunnels which were subsequently fixed in 2.5% glutaraldehyde in 0.1 mol/L phosphate buffer and stored at 4º C prior to use. Samples were then dehydrated using a graded ethanol series and dried using a critical-point dryer (Balzars 030 critical-point dryer, BAL-TEC AG, Furstetum, Liechtenstein). Dried samples were then mounted on stubs and coated with gold-palladium using a Pelco Model 3 sputter-coater (Ted Pella, Inc.). All samples were visualized with a Hitachi S-4700 cold field emission SEM (Hitachi Inc., Pleasanton, California) and all sample processing took place at the University of Montana EMtrix facility.



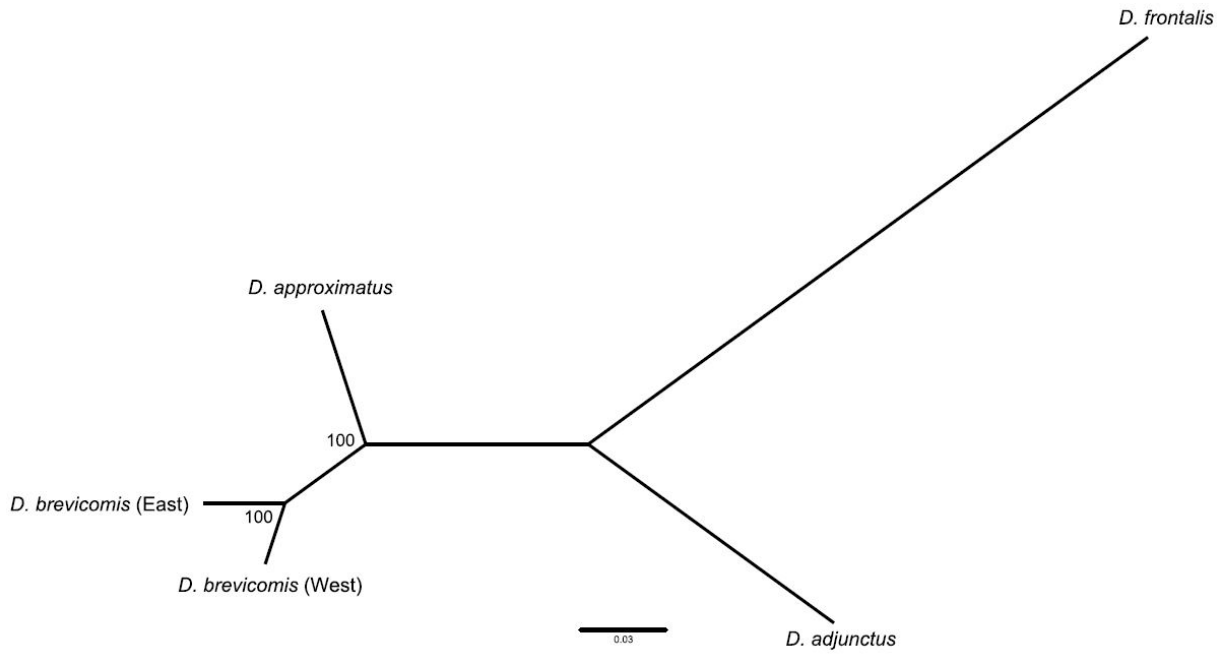
Supplemental Figure 1. Symbiotic fungi of the western pine beetle. Scanning electron microscope (SEM) images of a western pine beetle larval tunnel (with larva removed) and symbiotic fungal growth. The top left image shows the larval tunnel (dashed outline) and the location of subsequent magnified images (1-5).



Supplemental Figure 2. Phylogenetic networks of the basidiomycete and ascomycete. SplitsTree networks for the basidiomycete (top) and ascomycete (bottom) isolates collected from western pine beetles in the western U.S.



Supplemental Figure 3. Identifying X-linked scaffolds in the western pine beetle genome assembly. Shown is the normalized sequencing coverage of male and female RADtags over scaffolds in the western pine beetle genome build. Each scaffold shown as a point. Green points are scaffolds with significant BLAST hits to putative X chromosome scaffolds in the mountain pine beetle. Scaffolds below the hashed grey line (female coverage >1.5× male coverage) were considered putative X-linked scaffolds.



Supplemental Figure 4. Phylogenetic relationships among closely related prothoracic mycangium-bearing *Dendroctonus* species. Unrooted maximum likelihood (RAxML, GTR + gamma) tree of concatenated SNVs (33,808) from East and West western pine beetles and three close relatives.

Supplemental Table 1. Genome assembly statistics for the beetle (*Dendroctonus brevicomis*) and the basidiomycete (*Entomocorticium* sp. B).

Organism	Number of scaffolds	Total length with gaps	N50 contig length (kb)	N50 scaffold length (kb)	Maximum scaffold length	GC content	Estimated genome size	Estimated repetitiveness of genome
beetle (*Dendroctonus brevicomis*)	35,469	129,552,754	5	5	540,064	37%	191,539,754	28%
basidiomycete (*Entomocorticium sp. B*)	1,248	35,715,059	37	54	302,230	50%	64,149,417	54%

Supplemental Table 2. Results from BUSCO assessments of genome completeness for the basidiomycete, *Entomocorticium* sp. B and a closely related *Peniophora* sp. [11], and the beetle, *Dendroctonus brevicomis* and close relative *D. ponderosae* [10].

	Basidiomycete		**Beetle**	
	E. sp. B	*P*. sp. v1.0 (Ricme1)	*D. brevicomis*	*D. ponderosae*
Complete BUSCOs	267	271	1685	2016
Complete and single-copy BUSCOs	263	262	1516	1848
Complete and duplicated BUSCOs	4	9	169	168
Fragmented BUSCOs	15	10	619	518
Missing BUSCOs	8	9	371	141
Total BUSCO groups searched	290	290	2675	2675

Supplemental Table 3. Summary of collection locations and number of individuals genotyped per area. Beetles were collected using funnel traps baited with attractants (funnel), were extracted from trees (new attack) or reared out of bark from infested trees. For each species, per location, the type of genetic data used is presented and the number of individuals from the site is shown in parentheses.

Beetle collector	Beetle collection method	Location	Nearest town	Identifier	*E*. sp B	*C. brevicomi*	*D. brevicomis*	*P. ponderosa**
Ryan Bracewell	funnel	33° 28' N, 109° 22' W	Morenci, AZ	AP	RADseq(2)	NA	RADseq(5)	microsatellite(70)
Glenn Kohler	reared from bark	47° 16' N, 117° 34' W	Cheney, WA	CH	RADseq(2)	RADseq(1)	RADseq(5)	NA
Ryan Bracewell	funnel	42° 46' N, 121° 44' W	Chiloquin, OR	CQ	RADseq(2)	RADseq(1)	RADseq(4)	microsatellite(30)
Ryan Bracewell	reared from bark	46° 04' N, 114° 14' W	Darby, MT	DA	RADseq(1)	RADseq(1)	RADseq(5)	microsatellite(30)
Darci Carlson	funnel	47° 45' N, 120° 25' W	Entiat, WA	EN	RADseq(1)	NA	RADseq(5)	microsatellite(30)
Rich Hofstetter	funnel	35° 10' N, 111° 45' W	Flagstaff, AZ	FL	RADseq(2)	RADseq(1)	RADseq(8)	microsatellite(29)
Andy Graves	funnel	32° 53' N, 107° 45' W	Gila, NM	GI	RADseq(1)	NA	RADseq(3)	NA
Ryan Bracewell	funnel	45° 54' N, 120° 42' W	Goldendale, WA	GO	RADseq(1)	RADseq(1)	RADseq(5)	NA
Ryan Bracewell	funnel	48° 34' N, 118° 18' W	Kettle Falls, WA	KE	RADseq(1)	RADseq(1)	RADseq(5)	NA
Lia Spiegel	funnel	45° 19' N, 118° 19' W	La Grande, OR	LA	RADseq(1)	RADseq(1)	RADseq(5)	microsatellite(30)
Diana Six	funnel	46° 53' N, 113° 28' W	Greenough, MT	LF	RADseq(2)	RADseq(1)	RADseq(5)	microsatellite(26)
Ryan Bracewell	reared from bark	44° 06' N, 115° 21' W	Lowman, ID	LO	RADseq(1)	RADseq(1)	RADseq(5)	NA
Cynthia Snyder	funnel	41° 17' N, 122° 03' W	McCloud, CA	MC	RADseq(1)	NA	RADseq(5)	NA
Ryan Bracewell	funnel	46° 49' N, 114° 08' W	Missoula, MT	MI	RADseq(1)	RADseq(1)	RADseq(7)	NA
Ryan Bracewell	funnel	36° 18' N, 115° 40' W	Mt. Charleston, NV	MTC	RADseq(1)	RADseq(1)	RADseq(4)	microsatellite(30)
Danny Cluck	funnel	40° 41' N, 121° 13' W	Old Station, CA	OL	RADseq(1)	RADseq(1)	RADseq(4)	microsatellite(34)
Ryan Bracewell	funnel	37° 21' N, 107° 19' W	Piedra, CO	PA	RADseq(2)	NA	RADseq(8)	microsatellite(29)
Steve McKelvey	funnel	38° 44' N, 120° 44' W	Placerville, CA	PL	RADseq(1)	RADseq(1)	RADseq(5)	NA
Ryan Bracewell	reared from bark	44° 27' N, 120° 26' W	Prineville, OR	PV	RADseq(1)	RADseq(1)	RADseq(5)	NA
Ryan Bracewell	new attack	33° 28' N, 105° 44' W	Ruidoso, NM	RO	RADseq(1)	RADseq(1)	RADseq(8)	microsatellite(30)
Tom Coleman	funnel and bark	34° 10' N, 116° 55' W	San Bernardino, CA	SB	RADseq(2)	RADseq(1)	RADseq(7)	microsatellite(30)
Beverly Bulaon	funnel	37° 59' N, 120° 05' W	Tuolumne City, CA	SierraI	RADseq(1)	RADseq(1)	RADseq(2)	NA
Beverly Bulaon	funnel	38° 13' N, 120° 22' W	Avery, CA	SierraII	RADseq(1)	RADseq(1)	RADseq(5)	NA
Beverly Bulaon	funnel	37° 02' N, 119° 14' W	Bretz Mill, CA	SierraIII	RADseq(1)	NA	RADseq(5)	microsatellite(30)
Ryan Bracewell	funnel/bark	49° 41' N, 119° 46' W	Summerland, BC	SU	RADseq(1)	RADseq(1)	RADseq(8)	NA
Darci Carlson	funnel	48° 16' N, 120° 11' W	Twisp, WA	TW	RADseq(1)	RADseq(1)	RADseq(5)	NA
Ryan Bracewell	funnel	37° 35' N, 112° 15' W	Tropic, UT	UT	RADseq(1)	RADseq(1)	RADseq(8)	NA
Ryan Bracewell	reared from bark	44° 53' N, 115° 42' W	Yellow Pine, ID	YE	RADseq(1)	RADseq(1)	RADseq(5)	NA
Rich Hofstetter	funnel	34° 32' N, 112° 32' W	Prescott, AZ	PR	RADseq(1)	NA	RADseq(5)	NA

* Data from [12]

REFERENCES

[1] Gnerre, S., MacCallum, I., Przybylski, D., Ribeiro, F.J., Burton, J.N., Walker, B.J., Sharpe, T., Hall, G., Shea, T.P., Sykes, S., et al. 2011 High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl. Acad. Sci.* **108**, 1513-1518.
[2] Zimin, A.V., Marcais, G., Puiu, D., Roberts, M., Salzberg, S.L. & Yorke, J.A. 2013 The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669-2677.
[3] Leggett, R.M., Clavijo, B.J., Clissold, L., Clark, M.D. & Caccamo, M. 2014 NextClip: an analysis and read preparation tool for Nextera Long Mate Pair libraries. *Bioinformatics* **30**, 566-568.
[4] Etter, P.D., Bassham, S., Hohenlohe, P.A., Johnson, E.A. & Cresko, W.A. 2011 SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods in Molecular Biology* **772**, 157-178.
[5] Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. & Hoekstra, H.E. 2012 Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**, e37135.
[6] Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A. & Cresko, W.A. 2013 Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124-3140.
[7] Li, H. & Durbin, R. 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760.
[8] Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & Proc, G.P.D. 2009 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079.
[9] Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. 2011 The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158.
[10] Keeling, C.I., Yuen, M.M., Liao, N.Y., Roderick Docking, T., Chan, S.K., Taylor, G.A., Palmquist, D.L., Jackman, S.D., Nguyen, A., Li, M., et al. 2013 Draft genome of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins, a major forest pest. *Genome Biol.* **14**, 1-20.
[11] Nagy, L.G., Riley, R., Tritt, A., Adam, C., Daum, C., Floudas, D., Sun, H., Yadav, J.S., Pangilinan, J., Larsson, K.-H., et al. 2016 Comparative genomics of early-eiverging mushroom-forming fungi provides insights into the origins of lignocellulose decay capabilities. *Mol. Biol. Evol.* **33**, 959-970.
[12] Potter, K.M., Hipkins, V.D., Mahalovich, M.F. & Means, R.E. 2015 Nuclear genetic variation across the range of ponderosa pine (*Pinus ponderosa*): phylogeographic, taxonomic and conservation implications. *Tree Genet. Genomes* **11**, 1-23.