



The National Forest Genetics Laboratory

Use of genotyping technologies to identify tree variability

Valerie Hipkins, NFGEL Director

vhipkins@fs.fed.us



U.S. FOREST SERVICE

Caring for the land and serving people

Outline

- Ponderosa Pine (*Pinus ponderosa*)
 - ✓ Source ID
 - ✓ Building a database
 - ✓ Modeling unsampled areas
- Golden Chinquapin (*Chrysolepis chrysophylla*)
 - ✓ Assessing rangewide variation
 - ✓ Building database for management decisions to drive conservation
 - ✓ Marker choice
- Mulanje Cedar (*Widdringtonia whytei*)
 - ✓ Matching seed to origin
 - ✓ Marker choice
 - ✓ DNA isolation from wood

The National Forest Genetics Lab (NFGEL)

CONTEXT TO ALL PROJECTS

- **QUESTION TO BE ANSWERED**
 - Species
 - Scope (taxonomic, source, individual ID)
 - Application (what, where, when)
 - Use (access over time, # users)
- **AVAILABLE SAMPLE MATERIAL**
 - DNA Isolation (quantity and quality)
- **MARKERS (existence and appropriateness)**
- **REFERENCE MATERIAL (databases; maps)**
- **TIME (inquiry to answer)**
- **COST**
- **FACILITY/EXPERTISE**

Ponderosa Pine: Looking to it's Genes to Ensure it's Future

US Forest Service National Forest Genetics Lab, USFS Region 1, and Bureau of Land Management (BLM)

Ponderosa pine is vulnerable to climate shifts, wildfires, bark beetle attack, and loss of habitat due to development. Genetic diversity is an essential component of long-term forest health.



Local seed source

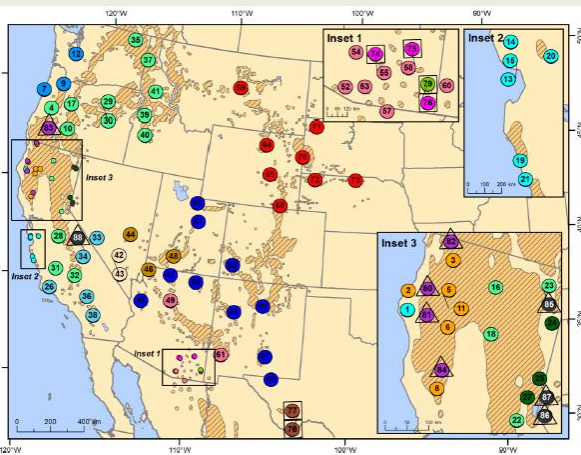
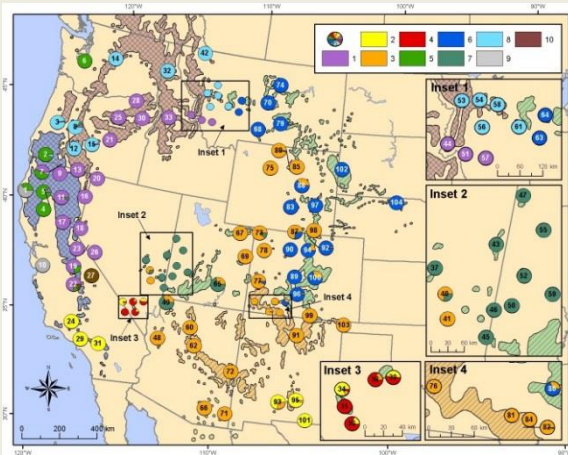


Offsite seed source



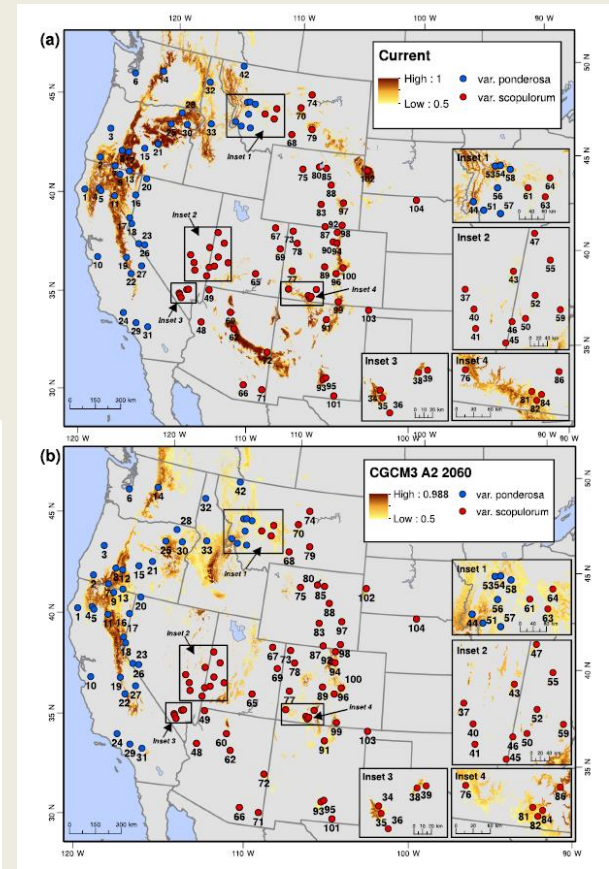
Identifying Original Seed Source of Plantations

- Reference Material
 - 114 stands over 3,000 trees; needle collections*
 - Samples received from May 2001 – Aug 2012*
- Markers
 - allozymes; microsatellites (nuclear and cp); mtDNA*
- Database Use
 - over 10-yr of lab work; \$550,000+ (Federal Gov \$)*
 - 4 peer-reviewed publications*
 - source ID of known stands*



Potter, Hipkins, Mahalovich, and Means. 2013.
AJB 100: 1562

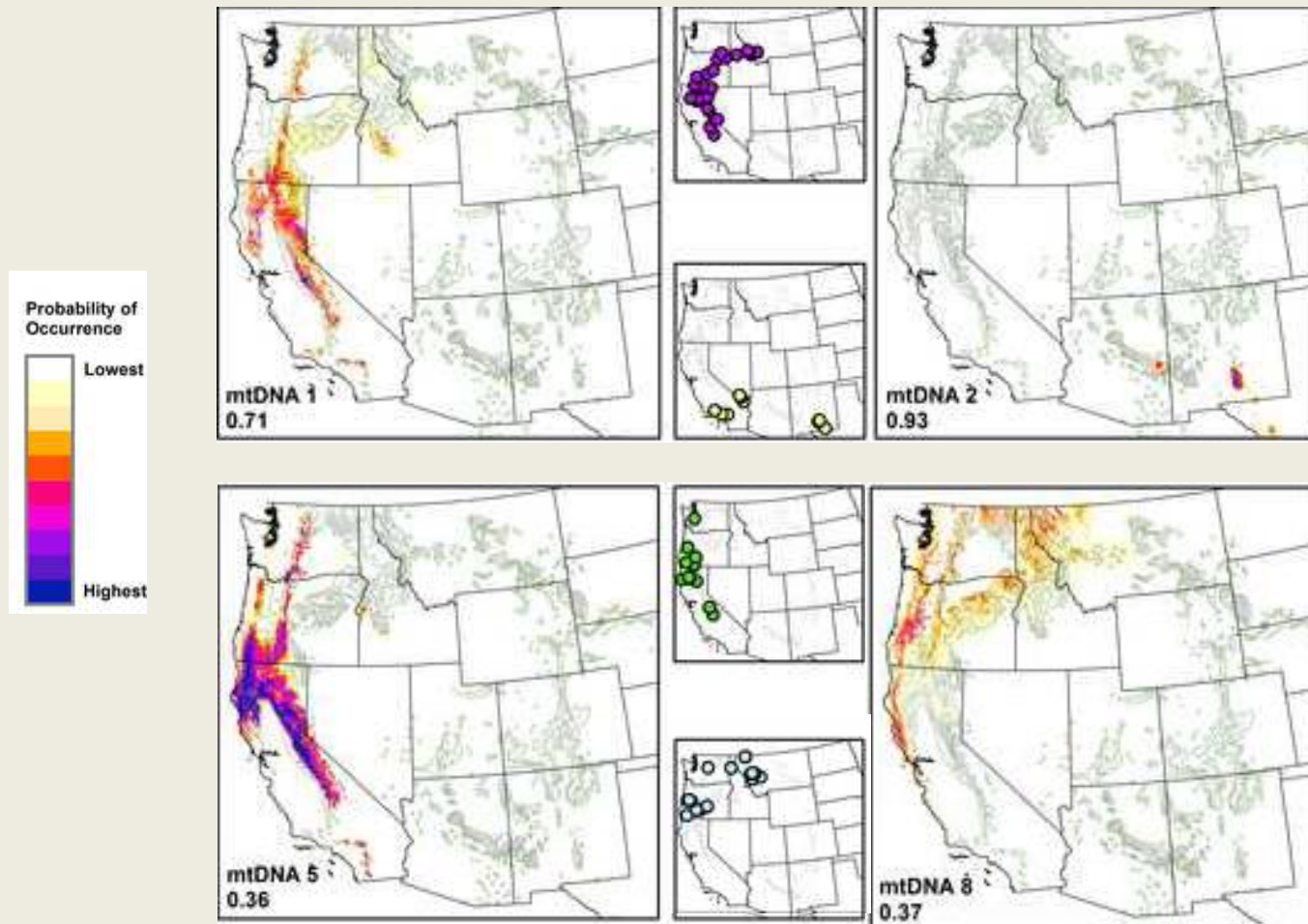
Willyard, Gernandt, Potter, Hipkins, Marquardt,
Mahalovich, Langer, Telewski, Cooper, Douglas,
Finch, Karemera, Lefler, Lea, and Wofford. 2017.
AJB 104: 161



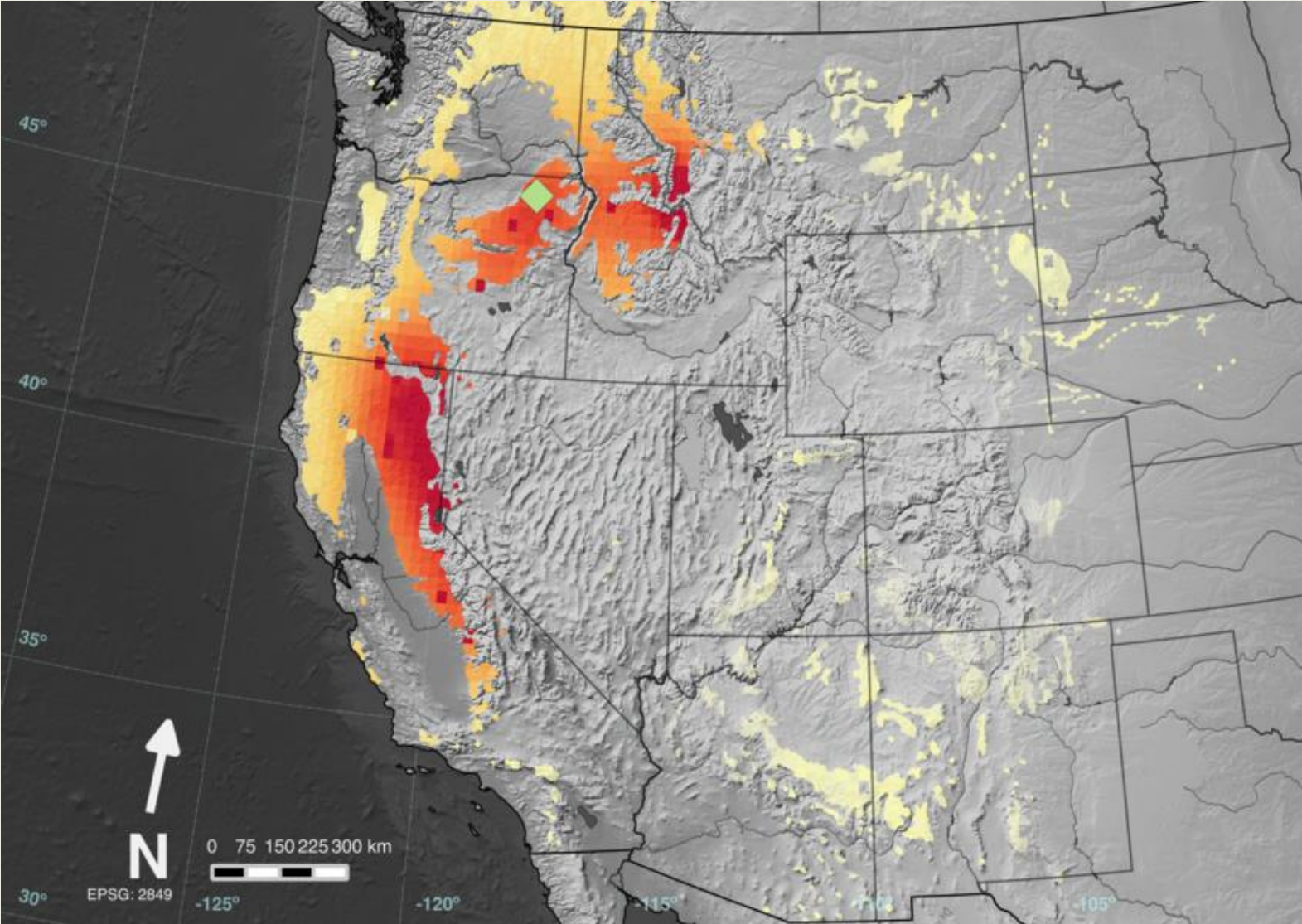
Potter, Hipkins, Mahalovich, and Means. 2015.
Tree Genetics & Genomes 11: 38

PONDEROSA PINE SOURCE ID

Probability of Occurrence



Relative likelihood of origin for the Round Mountain stand based on mitochondrial haplotypes.
Dr. Brook Milligan, Department of Biology, New Mexico State University



Target-Enriched High-Throughput Sequencing

(Ann Willyard, Hendrix College, USA)

PINUS GENOMES

Nuclear:

- All diploids (n=12)
- Massive genomes: 28 pg (*Arabidopsis* 0.2 pg)
- High levels of intra-specific diversity
- Incomplete lineage sorting
- Draft genome (haploid DNA from seed megagametophyte)
 - *P. taeda*; sister to subsection *Ponderosae*

Organelles:

- Plastids paternally inherited → pollen vs. seed movement (cp and mtDNA)
- Ancient plastid transfer within taxonomic subsection
- Plastid Simple Sequence Repeats (SSRs)
- Mitochondrial haplotypes - complex minisatellite

Target-Enriched High-Throughput Sequencing

(Ann Willyard, Hendrix College, USA)

Use HTS so obtain several hundred gene trees to build good species tree.

Considered Low Copy Nuclear Genes:

- **11,396 low-copy expressed genes**
- **6 Sanger-sequenced for previous phylogeny**
- **3 conserved ortholog set**

Chose 1045 genes to target:

- **Longer than 850 bp**
- **96 - linkage mapped**

Prelim: nearly complete plastomes; alignments for 285 nuclear genes; alignments for 5 mitochondrial genes. Ambiguous calls; missing samples at some genes; messy/missing data.

Golden Chinquapin (*Chrysolepis chrysophylla*)



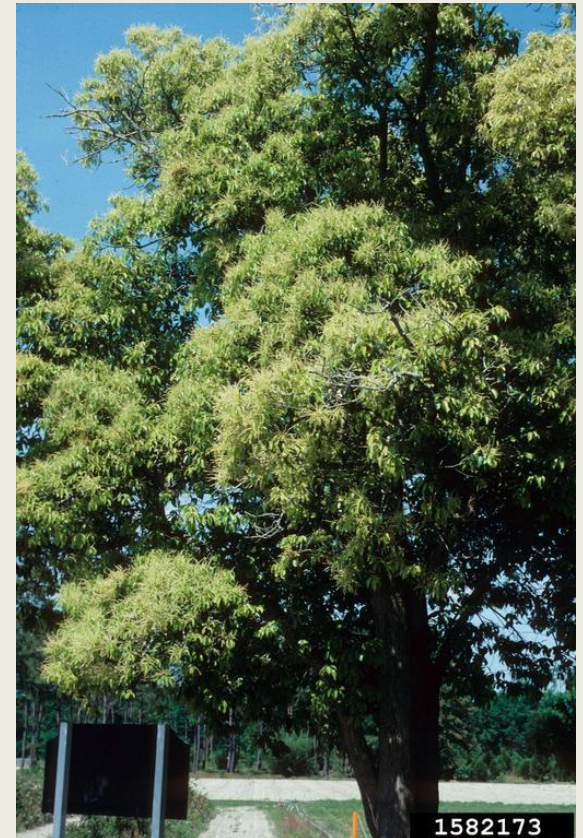
Pete Veilleux, East Bay Wilds. USFS

- **Beech family (Fagaceae)**
- **Endemic to western U.S.**
- **Evergreen tree about 20 – 40m in height (coastal, low elevation)**
- **Sensitive species in WA**

Objectives:

- **What is the level and distribution of genetic diversity within species**
- **How different are the northern-most disjunct populations from rest of range**
- **Are there varietal level genetic differences among the ‘shrub’ and ‘tree’ forms**

Project cooperators: Andy Bower, Ann Willyard, Eli Meyer, Jennifer DeWoody, Jacob Snelling



John Ruter, University of Georgia, Bugwood.org

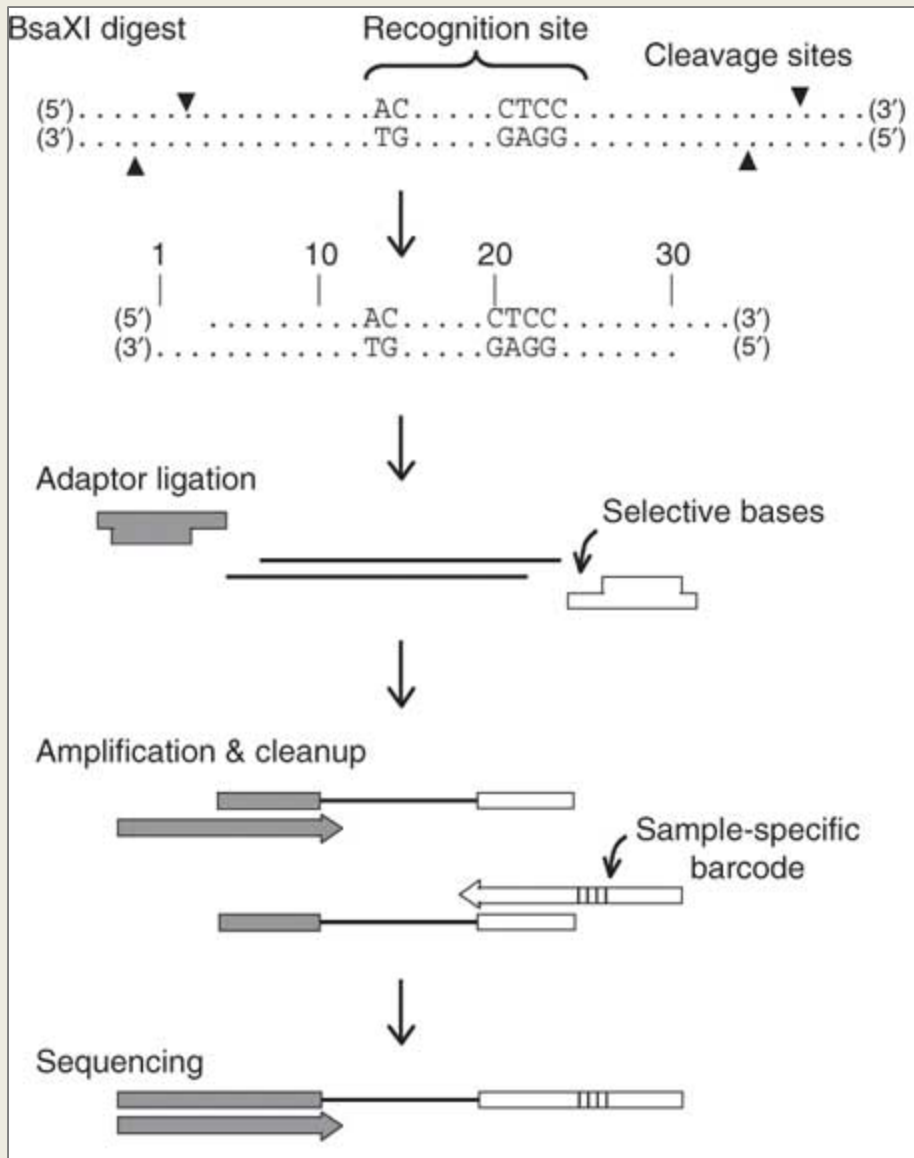
Marker Choice for Chinquapin

Allozymes: low variation obtained in tests

**Microsatellites (SSRs): pulled over from other Fagaceae
(oaks, European chestnut, etc)
16 loci, 716 trees, 23 stands
(~\$20,000; 1 yr)**

**2b-RAD: Next Generation Sequencing approach for
genome-wide genotyping (restriction site-
associated DNA (RAD), based on sequencing
fragments produced by type IIB restriction
endonucleases)
628 samples, 2021 SNPs with coverage \geq 20X
(~\$75,000; 2.5 yrs)**

SNP Genotyping Using 2b-RAD



Wang et al. (2012) Nat. Meth.

Genome reduction by type IIB RE

- Cuts in 2 sites, producing even-sized fragment with sticky ends
- Reduction using selective amplification (think AFLP)
- Produces equal-sized fragments for sequencing
- De novo reference construction based on overlapping tags

Comparison to RAD, GBS

- Flexible reduction (selection)
- Faster bench protocol
- *But* some questions about use in highly complex genomes

2b-RAD: Sequencing Results

Reads

Example Sequencing & Processing Stats (single lane):

Mean raw reads	1.63M per sample
Mean HQ reads	1.59M per sample
Mean mapped HQ reads	1.21M per sample
Mean N loci genotyped	96,351

SNP Matrix

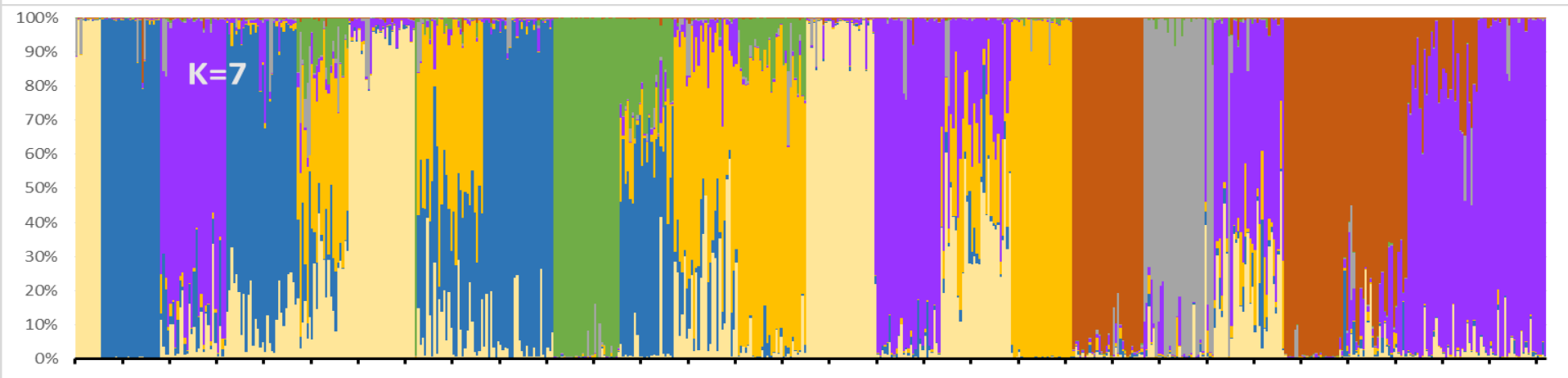
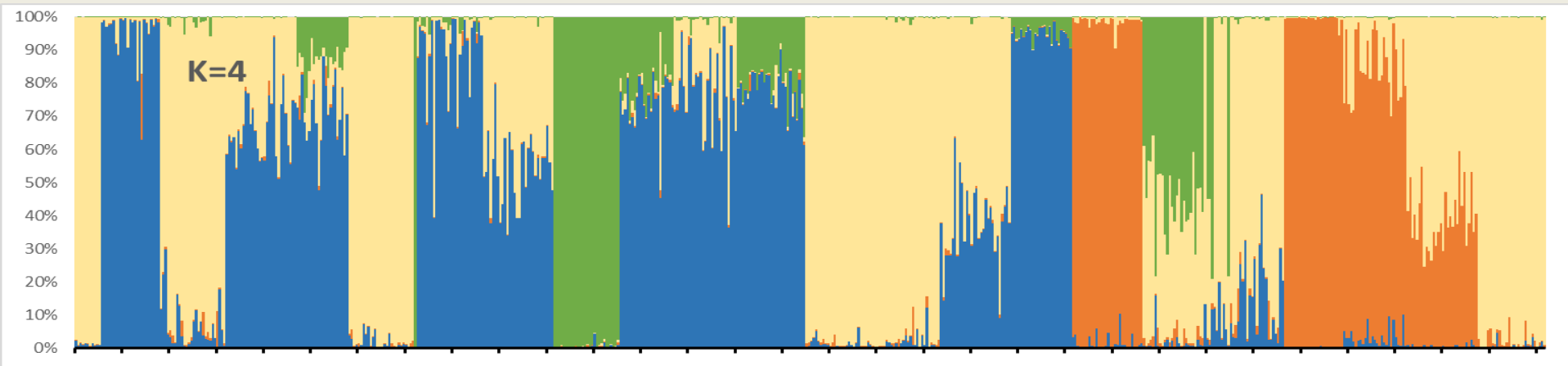
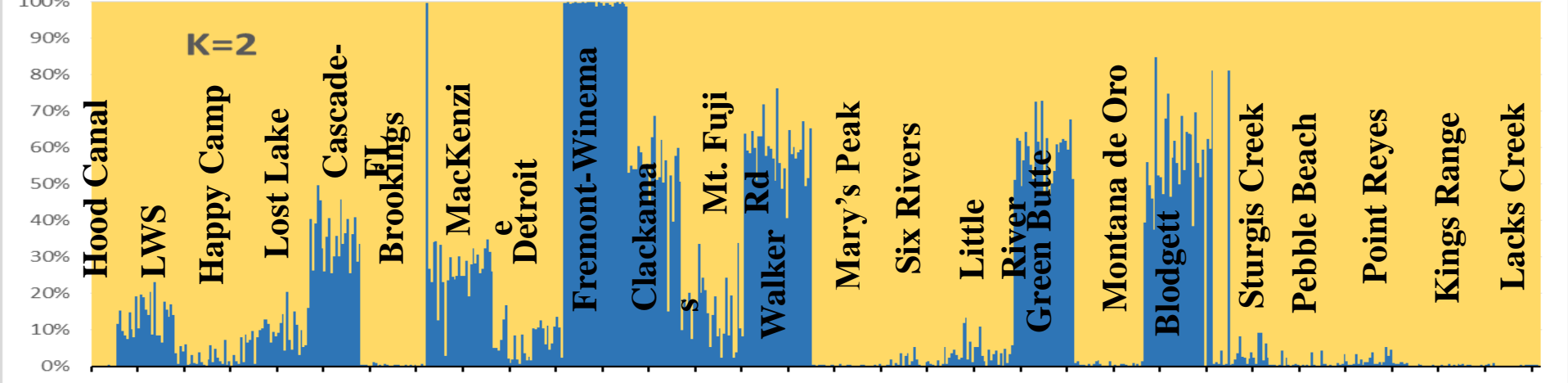
628 samples

2,021 SNPs with coverage $\geq 20X$

Mean 25% missing data per individual (range 10-62%)

SNP frequency over entire data set: 2.05%

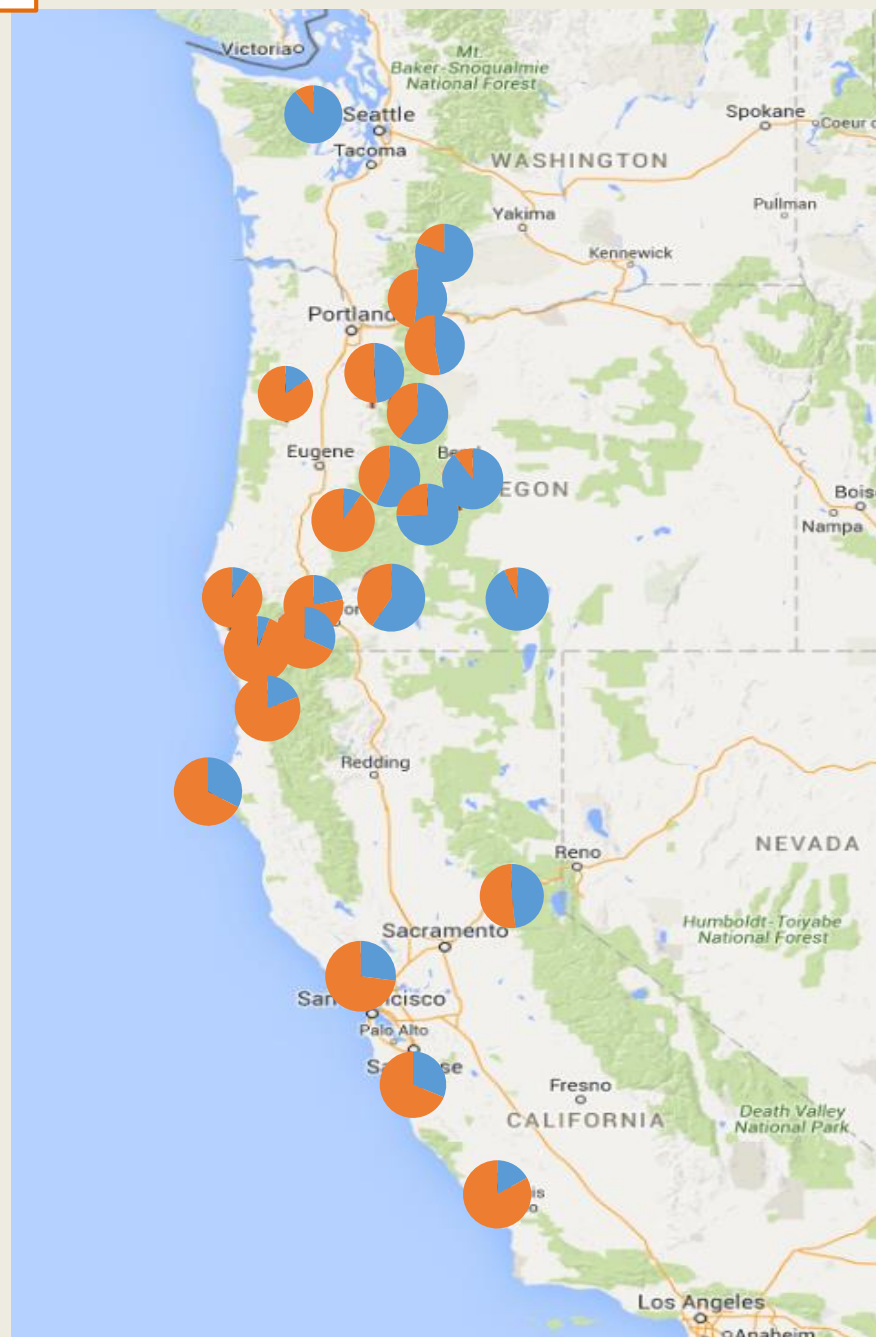
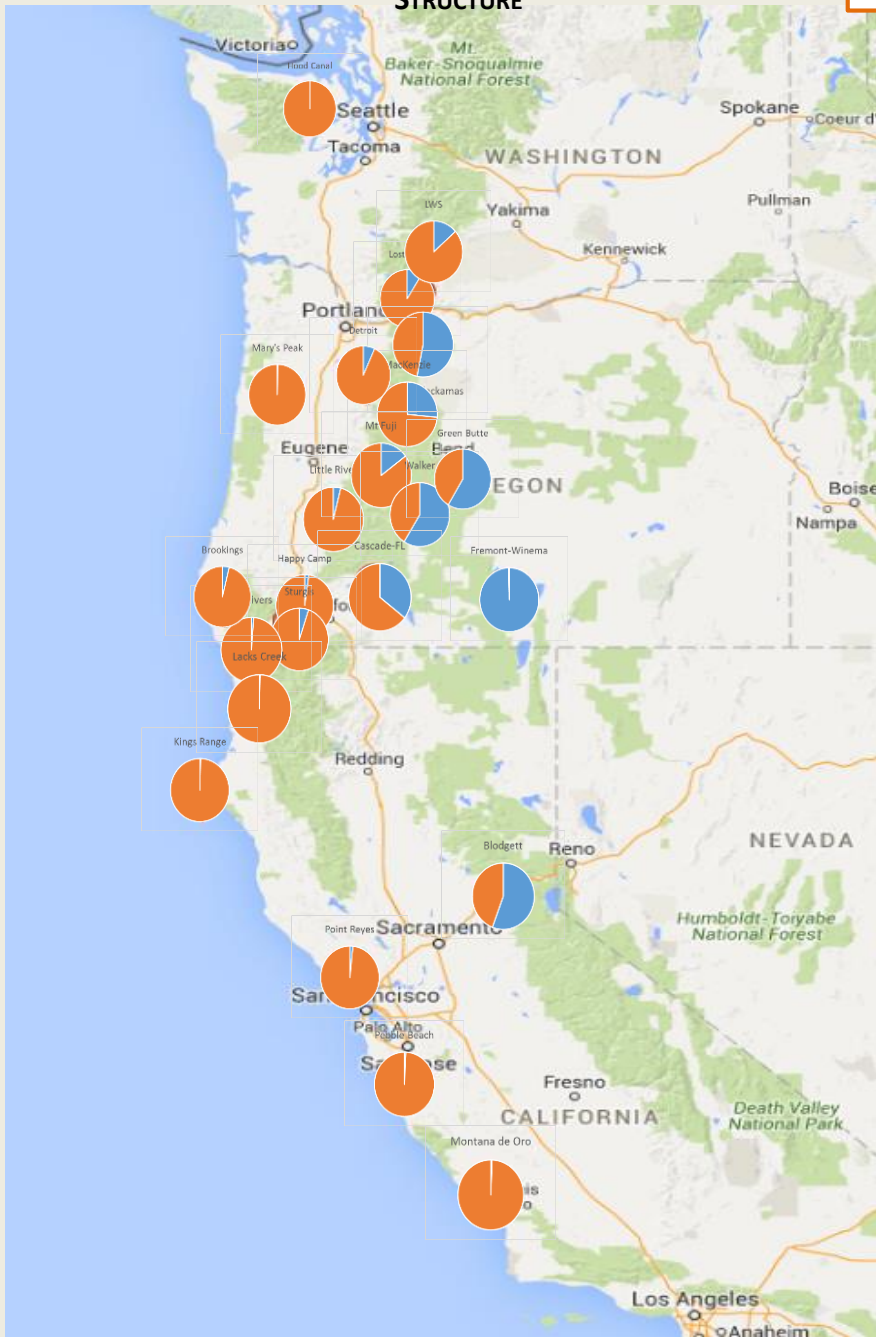
= 1 SNP every 48 bases



2b-RAD: Population STRUCTURE

K=2

SSR: Population STRUCTURE



Mulanje Cedar (*Widdringtonia whytei*)

Matching Seed to Origin

Background

- Malawi's national tree
- occurs naturally only in Mulanje Mtn Reserve
- critically endangered
- over-exploitation, theft, and fire
- discrete Regions/Basins/Stands
- 3 offsite stands producing seed; origin unknown
- match seed to appropriate planting source
- BGCI (Botanic Gardens Conservation International); Forestry Research Institute of Malawi; Mulanje Mountain Conservation Trust; USFS

Reference Database

- single native stand; plantations

Constraints

- inquiry: February 2017
- 700 wood samples, 150 des. tissue samples
- no markers
- final data requested by December 2017 (10 months)
- \$7,000

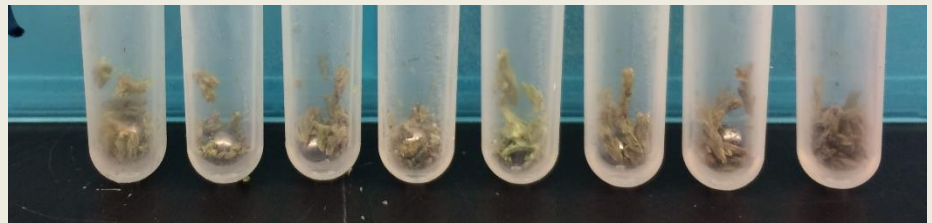
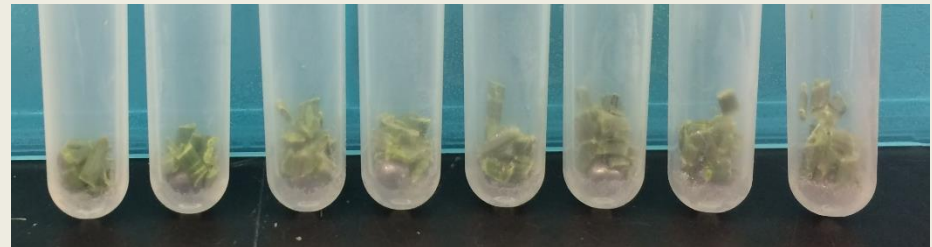


Mulanje Cedar (*Widdringtonia whytei*)

Matching Seed to Origin

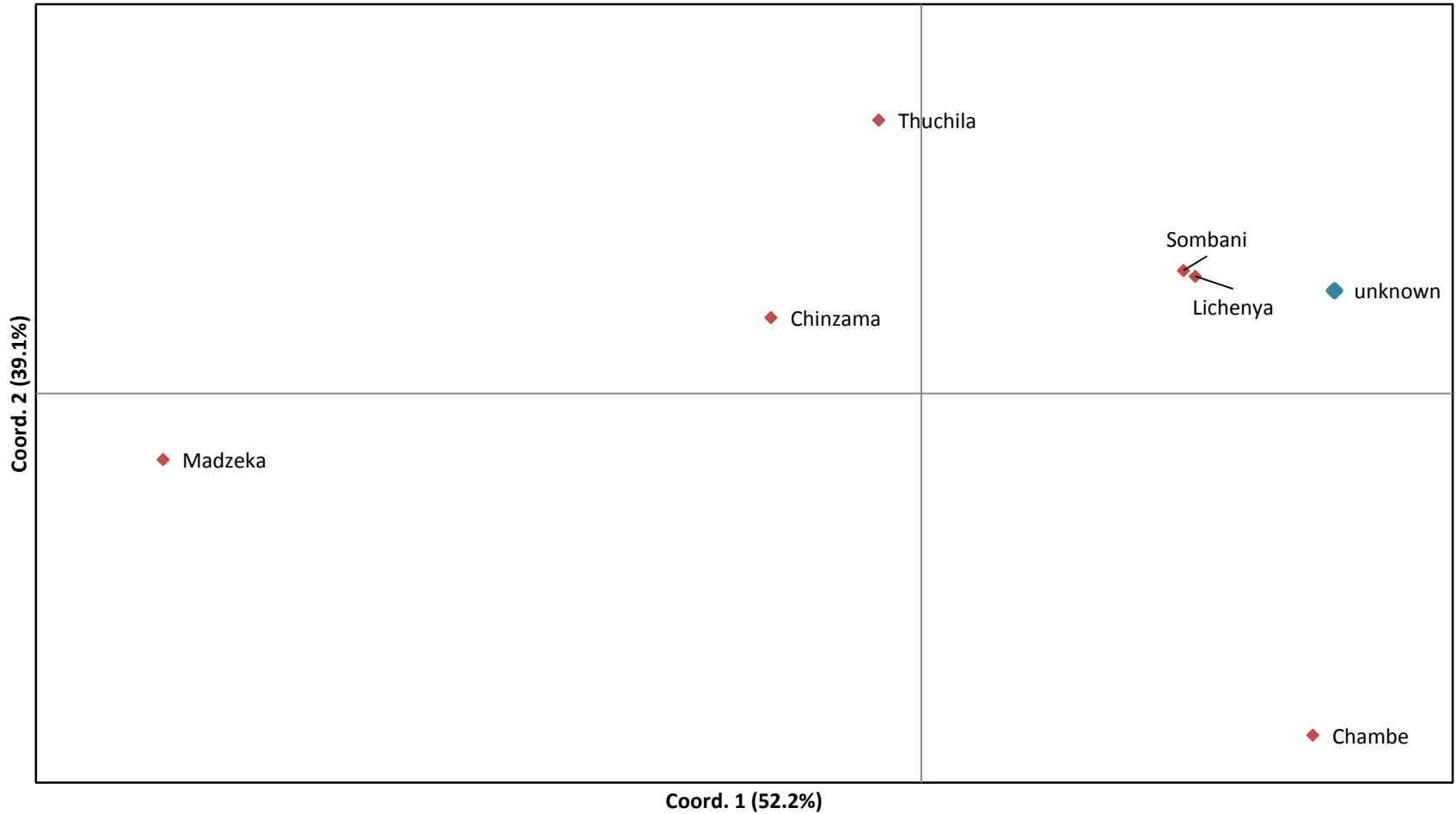
DNA Isolation and Markers

- 12 Microsatellites: “SSR markers from a thousand plant transcriptomes”. *Hodel et al. 2016. Applications in Plant Sciences 4:1600024; Hodel et al. 2016. Dryad Digital Repository. From Widdringtonia cedarbergensis.*



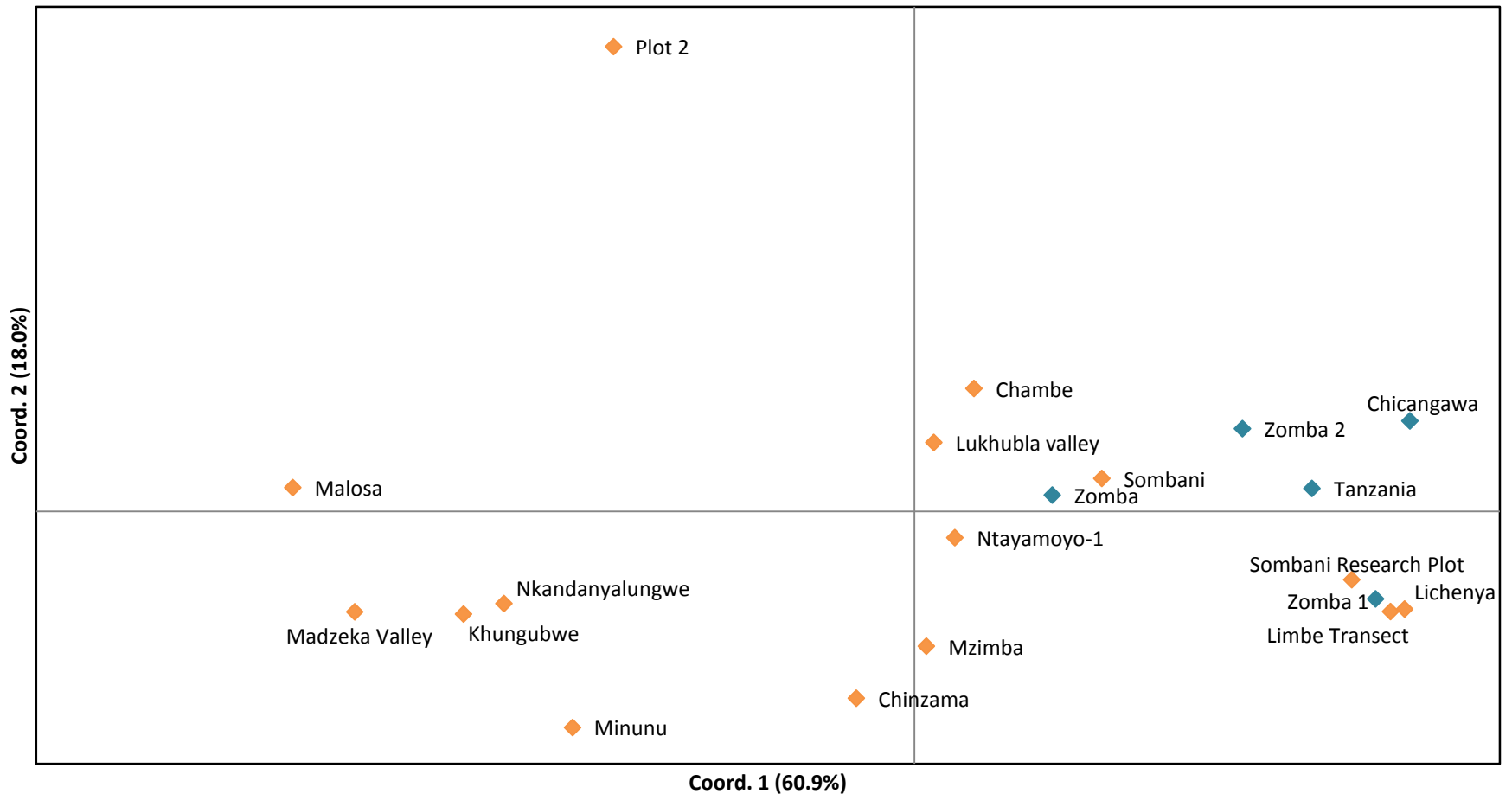
Mulanje Cedar (*Widdringtonia whytei*) Matching Seed to Origin

Principal Coordinates Analysis (PCA): 10 SSR loci; by basin; 190 trees



Mulanje Cedar (*Widdringtonia whytei*) Matching Seed to Origin

Principal Coordinates Analysis (PCA): 10 SSR loci; by stand; 190 trees



To identify tree variability.....

DEFINE FOCUS AND NEED
(species, scope, application)

REFERENCE DATABASES
(new technologies)

COOPERATION

