Tall Pinus luzmariae trees with genes from P. herrerae

Christian Wehenkel Corresp., 1, Samantha R Mariscal-Lucero 1, 2, M. Socorro González-Elizondo 3, Víctor A Aguirre-Galindo 1, Matthias Fladung 4, Carlos A López-Sánchez 5

1 Instituto de Silvicultura e Industria de la Madera, Universidad Juárez del Estado de Durango, Durango, Mexico
2 Instituto Tecnológico Valle del Guadiana, Durango, Mexico
3 CIIDIR Durango, Instituto Politécnico Nacional, Durango, Mexico
4 Thünen-Institute of Forest Genetics, Grosshansdorf, Germany
5 Escuela Politécnica de Mieres, Universidad de Oviedo, Mieres, Spain

Corresponding Author: Christian Wehenkel
Email address: wehenkel@ujed.mx

Context: Pinus herrerae and P. luzmariae are endemic to western Mexico, where they cover an area of more than 1 million hectares. Pinus herrerae is also cultivated in field trials in South Africa and South America, because of its considerable economic importance as a source of timber and resin. Seed quality, afforestation success and desirable traits may all be influenced by the presence of hybrid trees in seed stands. Aims: We aimed to determine the degree of hybridization between P. herrerae and P. luzmariae in seed stands of each species located in the Sierra Madre Occidental, Durango, Mexico. Methods: AFLP molecular markers from samples of 171 trees across five populations were analyzed with STRUCTURE and NewHybrids software to determine the degree of introgressive hybridization. Morphological analysis of 131 samples from two populations of P. herrerae and two populations of P. luzmariae was also conducted. The data were compared by Principal Coordinate Analysis (PCoA) in GenAlex 6.501. Results: Hybridization between Pinus herrerae and P. luzmariae was observed in all seed stands under study and resulted in enhancement of desirable silvicultural traits in the latter species. In P. luzmariae, all molecularly detected hybrids correspond to those identified on a morphological basis. However, the morphology of P. herrerae is not consistent with the molecularly identified hybrids from one population and is only consistent with 40% of those from the other population. Conclusions: This is the first report of hybrid vigour (heterosis) in Mexican pines. Information about hybridization and introgression is essential for developing effective future breeding programs, successful establishment of plantations and management of natural forest stands. Understanding how natural hybridization may influence the evolution and adaptation of pines to climate change is a cornerstone to sustainable forest management.
Tall *Pinus luzmariae* trees with genes from *P. herrerae*

Christian Wehenkel*1*, Samantha del R. Mariscal-Lucero2, M. Socorro González-Elizondo3, Víctor A. H. Aguirre-Galindo1, Matthias Fladung4, Carlos A. López-Sánchez5

1 Instituto de Silvicultura e Industria de la Madera, Universidad Juárez del Estado de Durango, México. E-mail: wehenkel@ujed.mx, vaguirre@mex.tuv.com

2 Instituto Tecnológico Valle del Guadiana, Durango, México. E-mail: crocio@ujed.mx

3 CIIDIR Durango, Instituto Politécnico Nacional, México. E-mail: herbario_ciidir@ipn.mx

4 Thünen-Institute of Forest Genetics, Germany. E-mail: matthias.fladung@thuenen.de

5 Escuela Politécnica de Mieres, Universidad de Oviedo, Mieres, Spain. E-mail: lpezscarlos@uniovi.es

**Key message:** The occurrence of hybrids in seed stands may influence seed quality and afforestation success. Here, we provide the first report of hybrid vigour in seed stands of the Mexican tree species *Pinus luzmariae* × *P. herrerae*.

**Abstract**

**Context:** *Pinus herrerae* and *P. luzmariae* are endemic to western Mexico, where they cover an area of more than 1 million hectares. *Pinus herrerae* is also cultivated in field trials in South Africa and South America, because of its considerable economic importance as a source of timber and resin. Seed quality, afforestation success and desirable traits may all be influenced by the presence of hybrid trees in seed stands.

**Aims:** We aimed to determine the degree of hybridization between *P. herrerae* and *P. luzmariae* in seed stands of each species located in the Sierra Madre Occidental, Durango, Mexico.

**Methods:** AFLP molecular markers from samples of 171 trees across five populations were analyzed with STRUCTURE and NewHybrids software to determine the degree of introgressive hybridization. Morphological analysis of 131 samples from two populations of *P. herrerae* and two populations of *P. luzmariae* was also conducted. The data were compared by Principal Coordinate Analysis (PCoA) in GenAlex 6.501.

**Results:** Hybridization between *Pinus herrerae* and *P. luzmariae* was observed in all seed stands under study and resulted in enhancement of desirable silvicultural traits in the latter species. In *P. luzmariae*, all molecularly detected hybrids correspond to those identified on a morphological
basis. However, the morphology of *P. herrerae* is not consistent with the molecularly identified hybrids from one population and is only consistent with 40% of those from the other population.

**Conclusions:** This is the first report of hybrid vigour (heterosis) in Mexican pines. Information about hybridization and introgression is essential for developing effective future breeding programs, successful establishment of plantations and management of natural forest stands. Understanding how natural hybridization may influence the evolution and adaptation of pines to climate change is a cornerstone to sustainable forest management.

**Key words:** AFLP; tree breeding; STRUCTURE; *Australes*; interspecific hybrids; tree species; timber; wood; monitoring

**Introduction**

Hybridization represents an important evolutionary force that can introduce much more new genetic material than is created by mutation events (Anderson, 1949; Wright, 1964). It can also act as an additional, perhaps more abundant, source of adaptive genetic variation than mutation (Grant & Grant 1994), by allowing gene flow and recombination (Abbott et al., 2013; Hipp, 2018). Furthermore, hybridization is one of the key sources of species formation and diversity, and many species may have originated by this route (Linder & Risenberg, 2004; Blanckaert & Bank, 2018), perhaps even as much as between 30% and 80% of all species (Wendel et al., 1991). On the other hand, increasing rates of hybridization may also lead to the extinction of unique populations or species because of unsuccessful reproductive efforts or introgression with a more common species (Rhymer & Simberloff, 1996; Blanckaert & Bank, 2018). In times of rapid ongoing climate change, hybridization may thus contribute to further extinctions, sometimes weakening reproductive isolation among species (Owens & Samuk, 2019) as well as supporting the development of novel segregating genotypes that will speed up adaptation to changes in climate (Chunco, 2014; Menon et al., 2019). Knowledge of hybridization has therefore deep practical reasons. Besides the adaptation issues, the presence of hybrid trees in seed stands “contaminates” the species’ gene pool and thus may influence seed quality and afforestation success (Arnold & Hodges, 1995; Rieseberg & Carney, 1998).

The process of hybridization incorporates alleles from one species into the gene pool of another (Harrison, 1993). Interactions between the environment and genetic structure can thus
lead to segregation of a novel taxon from parental types. Depending on the degree of
differentiation, hybrid offspring of two or more plants of different taxa are sometimes identified
as species, subspecies or variants (Futuyma, 1998; Tovar-Sánchez & Oyama, 2004).

Hybrids often display post-mating reproductive difficulties relative to their ancestors.
These difficulties include hybrid weakness, sterility and fitness breakdown (Rieserberg &
Carney, 1998). However, hybrids are not necessarily uniformly unfit. On the contrary, some
genotypic classes may be equally fit or even fitter than the parental taxa (Arnold & Hodges,
1995; Mabaso et al., 2019). The first hybrid generation ($F_1$) tends to exceed the parental
generation in vegetative vigour or robustness, in a condition also known as heterosis. However,
early hybrid generations such as $F_2$ and $F_3$ are often less vigorous and fertile than their parents
due to the break-up of adaptive gene arrangements (Rieserberg & Carney, 1998).

Studies involving hybridization are often based on morphological traits. However, the
phenotypic expression of characters of one taxon in another does not necessarily indicate
hybridization. Similar characters may occur in species because of phenotypic plasticity,
convergent evolution or simply because of a common ancestry, as Linder et al. (1998) observed
in wild sunflower. Furthermore, morphological characters yield limited information when the
parents and their hybrids are affected by environmental factors such as disease or drought stress,
generating a wide range of phenotypic variability. This problem is increased by subsequent
backcrossing of the hybrids to either parent species, resulting in morphological characters that
become more similar to those of the backcrossed parent species (Chen et al., 2004).

Use of molecular markers to detect interspecific hybridization is more effective than
verification by morphological, chemical or cytogenetic analysis, especially as access is available
to an almost unlimited number of molecular markers (Rieseberg et al., 1993; Alexandrov &
Karlov, 2018; Jasso-Martinez et al., 2018). Introgressive hybridization in many plant species has
been identified by molecular data (Rieseberg et al., 1993; Arnold, 1997; Delgado et al. 2007;
Kaplan & Fehrer, 2007; McVay et al., 2017). These markers have been useful for diagnosing $F_1$
and derived hybrid generations, evaluating levels of gene flow among species and reconstructing
phylogenetic relationships between hybridizing taxa and their close relatives (Rieseberg et al.,
1993).

Amplified Fragment Length Polymorphism (AFLP) markers have been successfully used
to detect introgressive hybridization in plants (Guo et al., 2005; Shasany et al., 2005; Koeber et
al., 2013), specifically in pines (Xu et al. 2008; Stewart et al., 2010; Vasilyeva & Semerikov, 2014; Ávila-Flores et al., 2016). AFLP markers include a more or less large number of polymorphic, di-allelic loci and can be developed relatively easily and at a relatively low cost, even for species about which no prior genetic information is available (Mueller & Wolfenbarger, 1999; Hardy et al., 2003; Paun & Schönswetter, 2012). Possible disadvantages of the AFLP technique such as compiling standardized patterns in a database for interlaboratory use and future reference can be avoided by using specific procedures as recommended by Savelkoul et al. (1999). However, AFLP as dominant marker does not allow identification of homologous alleles and thus scoring of homozygote and heterozygote states (Mueller & Wolfenbarger, 1999).

Interspecific hybridization is very common in natural stands of the genus Pinus (Critchfield, 1967; Quijada et al., 1997; Conkle & Critchfield, 1988; López Upton et al., 2001; Delgado et al., 2007; Ávila-Flores et al., 2016; Stacy et al., 2017; Vasilyeva & Goroshkevich, 2018; Mo et al., 2019), because of very weak reproductive barriers between pine species (Little & Righter, 1965; Garrett, 1979; Dungey, 2001); this could be generalized across conifers with similar divergence history (Menon et al., 2018). Interspecific F1 hybrids in this genus are highly viable and fertile (Critchfield 1975), which complicates taxonomic classification (Martínez, 1948; Lanner, 1974). Genetic diversity is often high in Pinus because of the usually large populations, cross-fertilization, high mutation rates and long-distance dispersion of pollen and sometimes seeds (Gernandt et al., 2011), as well as interspecific hybridization and introgression (Critchfield, 1967, 1975; Quijada et al., 1997; Conkle & Critchfield, 1988; Ledig, 1998; López-Upton et al., 2001; Ávila-Flores et al., 2016). In addition, understanding the phylogenetic relationships between closely related species of pines is also difficult due to retention of ancestral alleles (Delgado et al., 2007; Hernández-León et al., 2013; Ortiz-Martínez & Gernandt, 2016).

Moreover, North American hard pines in the subsection Australes share plastid DNA lineages due to introgressive hybridization or incomplete lineage sorting (Ortiz-Martínez & Gernandt, 2016).

Pinus herrerae Martínez and Pinus luzmariae Pérez de la Rosa belong to the subsection Australes, a monophyletic group including 29 pine species (Gernandt et al., 2005; Hernández-León et al., 2013). Herrera's pine (P. herrerae), also known as Pinus teocote var. herrerae (Martínez) Silba, is endemic to western Mexico, where it covers an area of about 1 million hectares (M ha) (Comisión Nacional Forestal, 2009) in mountain ranges between 16° and 28° N,
at elevations ranging from 1,100 m to 2,800 m (Dvorak et al., 2007; Wehenkel et al., 2015). The species is used to produce construction timber and resin (Martínez, 1948). It is also cultivated as an exotic in field trials in South Africa and South America because of its typically very tall, straight trunk (Dvorak et al., 2007). *Pinus luzmariae* (three-needled egg-cone pine), previously known as *Pinus oocarpa* var. *trifoliata* Martínez, was first recognized as a separate species by Pérez de la Rosa (1998). This small to medium-sized tree species is endemic to Mexico and it has been reported as covering an area of about 200,000 ha (Comisión Nacional Forestal, 2009). However, its distribution is not clear because it has been included in the very wide range of *P. oocarpa* Schiede ex Schltdl. The two largest populations are documented in the southern Sierra Madre Occidental covering about 1,000 hectares in south Durango and about 600 ha in northern Jalisco, respectively. Although no uses have been documented for *Pinus luzmariae*, it may be used as a source of timber, in the same way as *P. oocarpa*. The number of mature individuals of this species in its natural habitat is decreasing (Pérez de la Rosa & Farjon, 2013).

Both pine species grow in the Madrean-tropical subregion of the Sierra Madre Occidental, at lower elevations (<2,400 m). *Pinus herrerae* often is dominating in subhumid areas whereas *P. luzmariae* occupies sites with poor soils, although sometimes they grow together (González-Elizondo et al., 2012c, González-Elizondo at al., 2013). The ecological niches of these two species are clearly defined by soil pH and climate in the State of Durango (Mexico) (Wehenkel et al., 2015).

Population genetics studies of *P. herrerae* are scarce (see Wehenkel et al., 2015) and of *P. luzmariae* non-existent. Hybrids between these two species have not yet been reported so far. The aim of the present study was therefore to use AFLP molecular markers and morphological traits to determine the degree of hybridization between *P. herrerae* and *P. luzmariae* in seed stands of each species located in the Sierra Madre Occidental mountain system, Durango, Mexico. Although *P. herrerae* and *P. luzmariae* are morphologically very different (Perry, 1991; Farjon & Styles, 1997; García-Arévalo & González-Elizondo, 2003; Pérez de la Rosa & Vargas Amado, 2009), they are genetically closely related and can thus, theoretically, easily hybridize with each other (Dvorak et al., 2000; Ortiz-Martínez & Gernandt, 2016; Gernandt et al., 2018). In addition, we tested the possible *P. luzmariae* hybrid individuals for clues of possible hybrid vigour (heterosis). We aimed to unravel introgressive hybridization between *P. herrerae* and *P. luzmariae*, under the assumption that effective pollen flow has occurred between the two species.
Material and methods

Study sites

Samples were obtained from trees grown in three Pinus herraeae (PH) and two P. luzmariae (PL) seed stands located in the Sierra Madre Occidental, state of Durango (NW Mexico) (collection permit SEMARNAT SGPA/DGVS/003644/18). The three P. herraeae seed stands were (1) Ranchito (PH-R), (2) Manchón del Abies (PH-A) and (3) Ventana (PH-V). The P. luzmariae stands were (4) Laguna (PL-L) and (5) Tacuache (PL-T). All seed stands are uneven-aged and located in natural populations (Table 1).

The three PH seed stands grow on slightly acidic soil (pH 5.2 ± 0.4 (standard deviation [SD]), with H\(^+\) representing 27.4 ± 6.2 SD of total exchangeable cations) (Wehenkel et al., 2015). The Julian date of the last frost date in spring (S\(_{day}\)) was 118 (equivalent to April 28) ± 13 days SD. The elevation ranges between 2,318 and 2,511 m above sea level in the study area, with annual rainfall between 1,046 and 1,116 mm. The mean temperature varies from about 11 to 13°C. The PL stands are also situated on slightly acidic soils with pH 5.0 ± 0.4 SD and H\(^+\) representing 29.3 ± 5.6% SD of total exchangeable cations, although at lower elevations with an earlier S\(_{day}\) and higher temperatures. Their elevation varies from 1,960 to 2,140 m above sea level, S\(_{day}\) is 77 (equivalent to March 18 ± 13 days SD, with annual rainfall of between 1,107 and 1,139 mm. The mean temperature ranges between 14 and 16°C.

The PH-R and PH-A stands are separated from PL-L and PL-T by a deep (1,400 m) canyon and by a distance of 8.1-11 km (Fig. 1). The three PH stands include typical specimens of P. herraeae, i.e. tall trees, of height up to 40 m. However, both populations of P. luzmariae under study showed uncommon increased fitness relative to other populations of the same species (e.g. Pérez de la Rosa, 1998; García-Arévalo & González-Elizondo, 2003), as they are taller (see more in Discussion).

Fluorescence-based semi-automated AFLP analysis

Needles were collected from a total of 171 adult, dominant and superior putative phenotypes, i.e. plus trees, according to previously described selection criteria (Wehenkel et al., 2015), of both P. herraeae and P. luzmariae (33-35 per stand). Dendrometric variables were also
recorded in all seed stands, including coordinates, height ($H$) and diameter at breast height ($DBH$) of each sampled tree. The samples were placed in individual tubes with a few drops of ethanol and stored at -10 °C until DNA extraction.

DNA was extracted using the QIAGEN DNeasy96 plant kit, according to the steps described in the product manual. DNA fingerprints were obtained by amplified fragment length polymorphism (AFLP), according to a modification of the protocol of Vos et al. (1995), outlined by Ávila-Flores et al. (2016). The restriction enzymes used were EcoRI (selective primer: 5'-$\text{GACTCGGTACCAATTN}$N-3') and MseI (selective primer: 5'-GATGAGTCCTGAGTA$\text{AN}$NNN-3'). The primer combination E01/M03 (EcoRI-A/MseI-G) was used in the pre-AFLP amplification.

Selective amplification was carried out with the fluorescent-labelled (FAM) primer pair E35 (EcoRI-ACA-3) and M63+C (MseI-GAAC). All PCR reactions were carried out in a Peltier Thermal Cycler (MJ Research, Waltham, Massachusetts, USA). The amplified restriction products were electrophoretically separated in a Genetic Analyzer (ABI 3100 16 capillaries), with a GeneScan 500 ROX internal size standard (Applied Biosystems, Foster City, California, USA). The size of the AFLP fragments was resolved with the GeneScan® 3.7 and Genotyper® 3.7 software packages (Applied Biosystems, Foster City, California, USA).

The amplified restriction products were scored automatically. Only high quality fragments above the signal threshold of 50 (minimum peak height) (according to ABI manual) and with a maximum peak width of 1.0, minimum fragment size of 75 base pairs (bp), maximum fragment size of 450 bp and tolerance +/- bp of 0.4 were considered. Two fragments were only considered when the peak-peak distance between the two signals was at least 0.5 bp. The quality and reproducibility of the analysis were verified by inclusion of reference samples in each plate and independent repetition (replicate PCRs) of analysis at least 16 samples (i.e. a minimum of 16 randomly chosen individuals from each plate). In all replicates, the AFLP pattern was the same as in the first analysis (Simetá-Rodríguez et al., 2014; Ávila-Flores et al., 2016).

Two binary AFLP matrices were generated from the presence (code 1) or absence (code 0) at probable band positions (Table S1). The bands detected represented the presence of a dominant genetic variant (plus phenotype) with unknown mode of inheritance of this band position (detected fragment length) (Vuylsteke et al., 1999; Kraus, 2000). The absence of a band indicated the presence of only recessive genetic (allelic) variants at the given position (locus). To
minimize the rate of size homoplasy (Vekemans & Hardy, 2004; Caballero et al., 2008) and technical artefacts (Kraus, 2000), only the polymorphic loci (fragment lengths) with frequencies of occurrence of between 5 and 95% were selected for study (SanCristobal et al., 2006).

**Defining pure individuals and molecular identification of hybrids**

The trees PH-V4, PH-V49, PH-V52, PH-V64 and PH-V127, and PL-T28, PL-T31, PL-T37, PL-T103, PL-T130 were defined as individuals of "pure" *Pinus herrerae* (PH) and *P. luzmarias* (PL), respectively, (hereinafter called pure individuals or pure trees) identified by their genetic affiliation probability and by their morphological traits (see details below).

When PH or PL stands include common hybrid trees, they should possess a genome that is a combination of alleles derived from trees belonging to both species. These hybrids can be detected by genetic data obtained from molecular marker analysis (Xu et al., 2008; Ávila-Flores et al., 2016).

The resulting AFLP loci from the 171 tree samples were used to determine the degree of introgressive hybridization between PH and PL in the analysis, conducted with STRUCTURE version 2.1 (Pritchard et al., 2000; Falush et al., 2007) and NewHybrids version 1.1 Beta 3 software (Anderson & Thompson, 2002). Both software programs have been used to identify putative hybrids in *Pinus* with dominant markers such as AFLP (Xu et al., 2008; Ávila-Flores et al., 2016). The systematic Bayesian clustering approach applying Markov Chain Monte Carlo (MCMC) estimation, as implemented in STRUCTURE was used to test the affiliation of individuals to each species. The MCMC process started by randomly assigning individuals to a pre-determined population (group or species) number (*K*) (here *K* = 2, Fig. 2). Repeated many times in the burn-in process (burn-in period of 10,000 cycles), comprising 100,000 iterations, variant (allele) frequencies were estimated in each population and individuals re-assigned using those frequency estimates. In the course of the process, the convergence progressed toward reliable membership probabilities of individuals to a population (or species) (Porras-Hurtado et al., 2013).

If the probability of PH (or PL) affiliation of a putative PH (or PL) tree was less than 95% according to STRUCTURE, then this individual was recorded as a candidate hybrid. The affiliation probability was measured by the proportion of the dominant STRUCTURE populations in the studied stands (Table S2). Individuals were identified as first-generation (*F*₁)
hybrids when the probability of PH affiliation with a PL tree was in the range 48–52% (Xu et al., 2008; Ávila-Flores et al., 2016).

Use of the Markov chain Monte Carlo (MCMC) methodology and 100,000 sweeps after BurnIn (10,000 cycles), the NewHybrids 1.1 software (Anderson & Thompson, 2002) is suitable for the situation studied here, where only two diploid species appear to be hybridizing. By applying this software, Anderson (2008) showed that just ten AFLP were adequate to accurately separate parental and \( F_1 \) genotypes from later generation hybrid classes. A sample of \( M \) individuals, putative pure individuals as well as hybrid individuals, is obtained and genotyped at the \( L \) loci of codominant and dominant genetic markers. This software contemplates six genotype classifications (pure species 1, pure species 2, \( F_1 \) hybrids, \( F_2 \) hybrids, and the first backcross generation to pure species 1 or pure species 2) and estimates the probability that each individual belongs to the different classes (Anderson & Thompson, 2002; Anderson, 2008; Xu et al., 2008) (Table S3). A tree was assigned to one of the hybrid classes with a posterior probability of at least 95%.

To visualize individual and species differences, Principal Coordinate Analysis (PCoA) was also performed using the binary AFLP data matrix produced, Nei’s Genetic Distance (Nei, 1972, 1978) and GenAlex 6.501 software (Peakall & Smouse, 2012). The PCoA diagrams were elaborated with the first, second and third coordinate.

The accuracy of the software STRUCTURE and NewHybrids (burn-in period of 10,000 cycles, 100,000 iterations) in detecting hybrids was quantified using the computer program Hybridlab 1.0 (Nielsen et al., 2006). However, this software simulates intraspecific hybrids from population samples of co-dominant nuclear genetic markers, whereas the AFLP-technique can detect only dominant genetic markers. Here, the accuracy corresponds with the number of correctly identified individuals for a hybrid generation over the actual number of individuals assigned to that generation (Marie et al., 2011).

Assuming that the fixed band differences between PH and PL were homozygous (expected for fixed polymorphisms), a subset of 11 diagnostic AFLP loci of the five pure PH and PL trees distinguishing the two parental species (100% of one reference parental species had the band whereas 0% of the other parental species did not) were used to simulate three intraspecific hybrid generations (\( F_1 (N = 50) \) and \( F_2, F_1\text{PL}, F_1\text{PH}, F_2\text{PL}, F_2\text{PH}, F_1\text{PL-PL} \) and \( F_1\text{PH-PH} \) backcrosses; \( N = 125 \) for each). The majority of the used AFLP loci did not show fixed band
differences between PH and PL. Consequently, it was not possible to reliably identify the heterozygote or homozygote state by means of the AFLP bands, as found also by LaRue et al. (2013). Nevertheless, these three simulated intraspecific hybrid generations were also created when using all polymorphic AFLP loci assuming that the AFLP band was always the dominant homozygote and the recessive variant the recessive homozygote. Finally, we performed STRUCTURE and NewHybrids analyses to estimate their accuracy using the two simulation datasets.

**Morphological detection of hybrids**

To test the results of the AFLP analysis, morphological analysis was conducted on samples from the same trees in populations PH-A, PH-V, PL-L and PL-T sampled for the AFLP analysis. At least 31 individual trees were analysed for cone traits and 11 for needle traits per species. Samples of branchlets, needles and cones were collected for taxonomic determination and morphometric examination, and voucher specimens were deposited in the CIIDIR herbarium (acronym according to Thiers, 2019), the collection of the Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional of the Instituto Politécnico Nacional (Tables S4 and S5). Morphological characters were selected from those used by Martínez (1948) and Pérez de la Rosa (1998) in their descriptions of _P. herrerae_ and _P. luzmariae_, respectively, as well as those used by García-Arévalo & González-Elizondo (2003) to distinguish these two species in the study area, considering sheath, needle and cone characters (Table 2, Tables S6 and S7, Fig. S1). Characters that do not possess discrete or different states between _P. herrerae_ and _P. luzmariae_ according to these authors were excluded from the analysis as they have no informative value for this study, e.g. number of leaves per fascicle (3 in both species), persistence of fascicle sheaths (persistent in both), and cone peduncle (peduncle present, oblique and about the same diameter in both species). Some of the characters were measured at x40 with the aid of a Carl Zeiss Discovery.V8 stereo microscope. Individual and species differences were pictured by PCoA.

In order to detect PH and PL hybrids identified by morphological traits, we first identified five pure PH (PH-V 4, 49, 52, 64, 127) and five pure PL (PL-T 28, 31, 37, 103, 130) trees, applying a genetic affiliation probability larger than 0.99 according to STRUCTURE and NewHybrids and clearly assignable by morphological traits. Since not every independent
morphological trait was normally distributed and continuous, the species assignation of each tree
and “morphological” hybrids was established by Random Forest (Breiman, 2001) using the caret
package and function “train” (Kuhn, 2008; Williams et al., 2018) available in the free statistical
application R 3.5.2 (R Development Core Team, 2018). For this purpose, the PH trees were
labeled with the value “1” (corresponding to the presence of PH) and the PL trees were labelled
with “0” (corresponding to the absence of PH) in this presence–absence classification model.
The model for both the cone traits and needle traits, respectively, was fit using a 5-fold cross-
validation repeated 10 times (i.e., using 80% of the dataset as training set and the remaining 20%
as testing set). Random Forest is a nonparametric tree-based classifier and hence does not require
variable scaling and can successfully handle non-normality (Strobl, Malley & Tutz, 2009) as
well as categorical and confounding variables (Dormann et al., 2013). Caret package (short for
Classification And REgression Training) is a complete framework for building machine learning
models (http://caret.r-forge.r-project.org).
Using all morphological traits listed in Table 2, a tree was classified as a possible PH (or PL)
tree if it was assigned to each of those species with the highest assignment probability (i.e., >
50%). If the posterior probability of PH (or PL) affiliation of a possible PH (or PL) tree was less
than 95% according to Random Forest, then, that tree was considered as a putative PH (or PL)
ybrid.
The predictive ability of the Random Forest model was evaluated using the True Skill Statistic
(TSS; Allouche, Tsoar & Kadmon, 2006) using the caret package in R (for more details see
Escobar-Flores et al., 2018). TSS (also known as the Hanssen–Kuipers discriminant) is an
appropriate alternative to Area Under a Receiver Operating Characteristic (ROC) Curve (AUC;
Fawcett, 2006) in cases where model predictions are formulated as presence–absence models and
an improvement to the widely used kappa. TSS not only accounts for both omission and
commission errors, but is not affected by the sample size of each class. The TSS is defined as
sensitivity + specificity – 1, and ranges from −1 to +1, where +1 indicates a perfect classification
model and values of zero or less indicate performance no better than random (Allouche, Tsoar &
Kadmon, 2006; Tatler, Cassey & Prowse, 2018).

Results
Molecular detection of hybrids
The AFLP primer combination yielded 348 polymorphic bands of 75-450 base pairs across all individual specimens of *Pinus herrerae* (PH) and *Pinus luzmariae* (PL). PH yielded 338 and PL 316 polymorphic bands. Both species shared 304 AFLP fragments (87.4% of polymorphic bands detected).

Fig. 2 shows the percentage of hybridization obtained with the STRUCTURE software, for $K = 2$, the three PH seed stands have a dominant genetic variant (blue) and the two PL seed stands contain another dominant genetic variant (red). Based on a 5% probability of introgression of gene content, 92 (53.8%) putative hybrids between PH and PL were found in all the seed stands analysed. Thus, 18% of the individuals in the Ranchito PH stand (PH-R) were putative hybrids; all PH individuals in the Manchón del Abies PH stand (PH-A) displayed genetic introgression from PL, and 14 of the 35 individuals in the Ventana PH stand (PH-V) were putative hybrids (40.0%). Regarding *P. luzmariae*, 30 (85.7%) putative hybrids were detected in the Laguna stand (PL-L), whereas seven individuals (21.2%) in the Tacuache *P. luzmariae* stand (PL-T) displayed introgression with *P. herrerae*. Five trees were first-generation hybrids ($F_1$), as indicated by introgression of between 45 and 55%; three trees in the PH-A stand and another two in the PL-L stand (Fig. 2, Table 3).

NewHybrids software clearly identified 65 (38%) putative hybrids between PH and PL (Table 3). No putative hybrids were found in the Ranchito PH stand (PH-R). In total, 25.7% of the individuals in the PH-A were putative hybrids, and two of the 35 individuals in the Ventana PH stand (PH-V) were putative hybrids (5.7%). A large majority (94.2%) of the individuals in the Laguna *P. luzmariae* stand (PL-L) were identified as putative hybrids, whereas 64% of the individuals in the Tacuache *P. luzmariae* seed stand (PL-T) displayed genetic introgression with *P. herrerae*. Only one tree, located in the PL-L stand, was detected as a first-generation hybrid ($F_1$) (Table 3).

The accuracy test showed that the method NewHybrids (NH) correctly assigned at least 88% of naturally occurring “pure” PH and PL individuals using 11 diagnostic AFLP and all 348 AFLP. STRUCTURE (STR) presented much more errors, especially with the 11 diagnostic AFLP. Using the 11 diagnostic loci, for both methods detections of 1st and 2nd ($F_2$, $F_1$PL and $F_1$PH backcrosses) generation hybrids were 100% and nearly 100%, for 3rd generation hybrids ($F_2$PL, $F_2$PH, $F_1$PL-PL and $F_1$PH-PH backcrosses) this decreased further to 0.59% in STR and 0.49% in NH (posterior probability (PP) of at least 95%). Using the all 348 AFLP, simulations
demonstrated lower rates of inaccurately than the test with 11 diagnostic loci. STR and NH correctly assigned 100% of simulated $F_1$ hybrids and nearly 100% of 2nd generation hybrids. Moreover, NH correctly assigned 100% of simulated 3rd generation hybrids, too. Using STR, a lower percentage of 3rd generation hybrids were correctly assigned (50%) (PP of at least 95%) (Table 4).

The results of the Principal Coordinates Analysis (PCoA) comparing genetic differences between individual specimens of PH and PL are shown in Fig. S2. At the individual level, the first three coordinates in PCoA explained 13.4% of the variability.

**Morphological detection of hybrids**

*Pinus herrerae* and *P. luzmariae* are morphologically distinct and easily recognized by several needle traits as well as by the width, scale position and scale length of the cone. However, various morphological intermediates between the two species were found (Fig. 3 and 4), including hybrids confirmed by Random Forest (Table 5, Fig. 5) considering seven cone (hybrid proportion of 4.6%) and eight needle traits (4.5%). Every observation was correctly classified (TSS = +1). At the individual level, the first three coordinates in PCoA only explained 36.2% of the variability in seven cone traits, but 61.7% of the variability in eight needle traits (Fig. S3 and S4). Hybrids identified by 15 morphological traits matched only 13.4% of the molecularly detected hybrids, and 5.7% of hybrids were only found by morphological traits.

**Clues of possible hybrid vigour (heterosis) in *P. luzmariae***

In this study of 69 *P. luzmariae* (pure and hybrid) trees, the hybrid heights and DBHs were much heterogeneously distributed than the dimensions of the pure trees. The smallest (one tree with 14 m height) and the tallest trees (14 trees with 23 - 30 m) were hybrids. The pure trees presented a normal distribution (probability) in which the expected proportion of trees higher than 24 m was much lower than the observed frequency of the tallest hybrids (Fig. 6).

**Discussion**

Species crossability in pines is of great theoretical and practical interest (Lopez et al., 2018; Vasilyeva & Goroshkevich, 2018). Many pine hybrids, including several Mexican species,
have been planted in trials across southern Africa in different conditions and climate regimes (Hongwane et al., 2018). Here, we report for the first time about the occurrence of hybrids in *Pinus luzmariae*, a little known species, introgressed by *P. herrerae*, revealing taller trees in comparison to all populations previously known for the species (as compared with those described in Pérez de la Rosa (1998) and García-Arévalo & González-Elizondo (2003)).

Populations of the introgressed *P. luzmariae* include trees 14 to 30 m (vs. 6-12 m in most other populations) (Fig. 6). This can be interpreted as hybrid vigour or heterosis, being the first report for Mexican pines. Other studies have shown that hybrid pines in the country do not differ from the pure trees in relation to vigour or robustness, e.g. *Pinus oocarpa × P. pringlei* (López-Upton et al., 2001) and *P. arizonica × P. engelmannii* (Ávila-Flores et al., 2016). In comparison with other populations of the same species, populations of the introgressed *P. luzmariae* (PL) display some important characters that are used to select superior forest trees (Kedharnath, 1984), i.e. good growth vigour, superior height, good self-pruning, and straight cylindrical bole. In the two introgressed populations, *P. luzmariae* formed almost pure, relatively dense stands with a few specimens of *P. herrerae* (PH) and *Pinus devoniana* in association, in contrast to the open, mixed stands in which *P. luzmariae* usually grows.

The high percentage of AFLP fragments (87%) shared by *P. herrerae* and *P. luzmariae* resulted in a large proportion of putative hybrids (54% by STRUCTURE (STR) and 38% by NewHybrids (NH)) using a posterior probability of at least 95%. The accuracy test detecting the different hybrid classes showed comparable results to other studies. However, NH detected more of both, accurate hybrids and individuals of pure species, than STR (Table 4). This explains the notable difference in the putative hybrid number found between STR (92 hybrids) and NH (65 hybrids). Therefore, the results presented by NH are probably more precise.

Previous studies have reported that the accuracy of these software can differ greatly depending on the population context (Vähä and Primmer 2006). Using AFLP and a PP of at least 90%, LaRue et al. (2013) presented 100% accuracy rate of simulated *F*₁ hybrids, but lower percentages of *F*₂ and backcrosses (about 91% and 92%, respectively). Cullingham et al. (2011) found a mean power of 74% to detect hybrids using microsatellites and a PP of at least 90% for *F*₁ hybrids and PP < 90% for other hybrid classes.

The high degree of introgression can be explained by the relatively recent diversification of species in the subsection *Australes* and the very weak reproductive barriers between them.
(Little & Righter, 1965; Garrett et al., 1979; Dvorak et al., 2000; Vargas-Mendoza et al., 2011; Gernandt et al., 2018). Similar weak reproductive barriers and high recent speciation rates have been recorded for madrones (Arbutus spp.) and oaks (Quercus spp.) (González-Elizondo et al., 2012a, 2012b, 2013; Hipp et al., 2019), the other two tree genera that, along with pines, dominate in the temperate forests of the Sierra Madre Occidental in western Mexico, where this study was carried out. Introgressive hybridization, although usually not obvious, may be more important in evolution than those cases in which hybridization is evident (Anderson, 1949; Hipp et al., 2019).

The interspecific gene transfer between the two pine species studied here is also supported by (i) wind pollination, (ii) longevity of individual trees, (iii) overlapping generations, (v) large effective population sizes, and (vi) weak physical barriers caused by sympatric distribution (Ávila-Flores et al., 2016). The relatively high diversity and high levels of gene flow in trees (relative to herbs and shrubs) is favoured by their outcrossed mating system and long distance seed dispersal (Petit & Hampe, 2006).

Of the five seed stands studied, PL-L displayed the highest degree of hybridization (94%), confirmed by AFLP as well as cone and needle traits (Tables 3 and 4). According to the PCoA results (Fig. S2, S3 and S4), many PL-L individuals were genetically closely related to P. herrerae individuals. The high phenotypic plasticity and more luxuriant growth found in both populations of P. luzmariae under study (Table 3 and 5) are a consequence of the hybrid origin. Crossbreeding or heterozygosity promotes variability, as found by Strauss (1987) for heterozygous trees of Pinus attenuata Lemm. derived from crossbreds. Resistance to disease, pathogens or environmental stresses has been a target in tree breeding towards interspecific hybrids. For example, Pinus patula Schltdl. & Cham. has been crossed with P. tecunumanii F.Schwerdtf. ex Eguiluz et J.P.Perry and with P. oocarpa in plantations in South Africa to increase tolerance to a fungal pathogen. The resulting hybrids of these three Mexican pines have a low frost tolerance, so new crosses were made until the finding that P. patula × P. tecunumanii from high elevations has a higher frost tolerance than P. patula × P. tecunumanii from lower elevations (Mabaso et al., 2019).

The high degrees of hybridization have several possible consequences: (i) extinction of one of the PH or PL parental species due to wasted mating effort or genetic swamping; (ii) reinforcement of species boundaries; (iii) creation of a third, hybrid species; (iv) formation of a
stable hybrid zone; and (v) partial introgression between the two hybridizing lineages (Chunco, 2014).

The stands of *P. luzmariae* we studied appear to represent a stable hybrid zone, like the *Pinus engelmannii* stand reported by Ávila-Flores et al. (2016). This conclusion is supported by the fact that the *P. luzmariae* population displays higher fitness than other populations of the species. Hybrid speciation does not occur in the studied populations as the hybrids are not spatially or ecologically isolated from the parental species, and no novel variants of morphological traits were found (Ungerer et al., 1998). The PH-A displayed the highest degree of hybridization in three PH studied (at least 37.1% consisting of at least 25.4% trees detected by NewHybrids and four extra trees detected by the cone and needle traits) (Tables 3 and 5). PH-A was located next to the two *P. luzmariae* seed stands (PL-L and PL-T), and it is expected that it intercepts large amounts of *P. luzmariae* pollen. We can, therefore, conclude that gene flow has occurred in both directions, from PL to PH and vice versa. However, gene flow from PH to PL seems to be more effective as more hybrids were found in the PL stands, both of which are located at lower elevations than the PH stands (Table 1).

Despite the large number of hybrids detected in the studied stands, the frequency of first-generation (*F*₁) hybrids and backcrossing was low (Table 3), indicating that hybrid crossing was usual in the seed stands. Similar results have recently been reported in species of *Salix* and for pines in the subsection *Ponderosae* and (Fogelqvist et al., 2015; Ávila-Flores et al., 2016).

Natural hybridization has also been observed in other Mexican pine species (Gernandt et al., 2018). Previous studies of the subsection *Australes* identified natural hybridization between *P. oocarpa × Pinus caribaea* and *P. oocarpa × Pinus pringlei* only by morphological traits (Styles et al., 1982; López-Upton et al., 2001). In a study of Mexican pine species of the subsection *Ponderosae*, Ávila-Flores et al. (2016) observed a high degree of introgressive hybridization between *P. engelmannii, Pinus arizonica, Pinus cooperi* and *Pinus durangensis*. AFLP markers detected most of the putative hybrids (58%), and only a few were detected by morphological features (15%). Hybridization was not detected by morphological traits in 74% of all hybrids detected by AFLP. Hybrids and backcrossing were also found in Mexican *Arbutus* species that are common in disturbed areas (González-Elizondo et al., 2012a; González-Elizondo et al., 2012b). Natural pairwise and triple hybrids have also been detected in numerous Mexican *Quercus* stands (e.g. Peñaloza-Ramírez et al., 2010; Hipp et al., 2019). Natural hybridization
between different *Populus* species and gene flow between cultivated poplars and native poplar populations have been described for European riparian forests and stands (e.g. vanden Broeck et al., 2004; Smulders et al., 2008; Lexer et al., 2005).

**Conclusions**

Hybridization between *Pinus herrerae* and *P. luzmariae* in seed stands in the Sierra Madre Occidental of Mexico has occurred in both directions to different degrees. Estimates of the success of hybrid individuals may be biased in this study by the fact that sampling was conducted in seed stands (in which plus trees predominate). Further research is necessary to increase our understanding of how hybridization may influence silvicultural traits in Mexican pines, as well as their evolution and adaptation to climate change. The successful survival and reproduction of these hybrids over generations will depend on their attributes, their fitness and the environmental factors influencing them (Strauss, 1987), given that hybridization leads to individuals which widely vary depending of the context, location and involved species (Gompert and Buerkle, 2016).

We conclude that both morphological and molecular approaches are essential to confirm the genetic identity of forest reproductive material as PH and PL frequently hybridize in all seed stands under study. Such information is very important for developing effective future breeding programs and successful establishment of plantations (Ávila-Flores et al., 2016) as well as for improving planning of the management of natural stands.

Introgressive hybridization in seed stands of *Pinus herrerae* and *Pinus luzmariae* generated outstanding plus trees. Because of their tall, straight trunks, hybrids of the largely unknown *Pinus luzmariae* represent a promising, valuable source of timber for wood industries as well as for reforestation in poor sites. The hybrid trees may be able to be cultivated after evaluation germplasm and vegetative propagation potential and may be suitable for commercial exploitation. However, further research is needed to examine the performance of hybrids and to assess their fertility and growth relative to those of pure species. Finally, monitoring natural hybridization is important in relation to sustainable forest management in Mexico.

**Acknowledgements**
We thank Sergio Simental-Rodriguez, Saskia Friedrich and Javier Hernández-Velazco for assistance with fieldwork and dataset preparation and Imelda Flores for assistance with the morphometric study.

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R, Butlin RK, Dieckmann U, Eroukhmanoff F, Grill A, Cahan SH, Hermansen JS, Hewitt G, Hudson AG, Jiggins C, Jones J, Keller B, Marczewski T, Mallet J, Martinez-Rodriguez P, Möst M, Mullen S, Nichols R, Nolte AW, Parisod C, Pfennig K, Rice AM, Ritchie MG, Seifert B, Smadja CM, Stelkens R, Szymbura JM, Väinölä R, Wolf JBW, Zinner D. 2013. Hybridization and speciation. J Evol Biol 26: 229-246. DOI: 10.1111/j.1420-9101.2012.02599.x

Alexandrov OS, Karlov GI. 2018. Development of 5S rDNA-based molecular markers for the identification of Populus deltoides Bartr. ex Marshall, Populus nigra L., and their hybrids. Forests 9:604 DOI: 10.3390/f9100604.

Allouche O, Tsoar A, Kadmon R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology 43(6):1223-1232.

Anderson E. 1949. Introgressive hybridization. Wiley. New York, USA.

Anderson EC, Thompson EA. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217-1229.

Anderson EC. 2008. Bayesian inference of species hybrids using multilocus dominant genetic markers. Philos Trans R Soc Lond B: B Sci 363(1505):2841-2850.

Arnold ML. 1997. Natural hybridization and evolution. Oxford University Press.

Arnold ML, Hodges SA. 1995. Are natural hybrids fit or unfit relative to their parents? Trends Ecol Evol 10:67-71.

Ávila-Flores IJ, Hernández-Díaz JC, González-Elizondo MS, Prieto-Ruíz JÁ, Wehenkel C. 2016. Degree of Hybridization in Seed Stands of Pinus engelmannii Carr. In the Sierra Madre Occidental, Durango, Mexico. PloS ONE 11(4):e0152651 DOI:10.1371/journal.pone.0152651.

Blanckaert A, Bank C. 2018. In search of the Goldilocks zone for hybrid speciation. PLoS Genet 14(9):e1007613 DOI: 10.1371/journal.pgen.1007613.
Bodenhofer U, Kothmeier A, Hochreiter S. 2011. APCluster: an R package for affinity propagation clustering. Bioinformatics 27:2463-2464.

Breiman L. 2001. Random forests. Machine Learn. 45:5-32.

Caballero A, Quesada H, Rolán-Alvarez E. 2008. Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. Genetics 179:539-554.

Chen J, Tauer CG, Bai B, Huang Y, Payton ME, Holley AG. 2004. Bidirectional introgression between Pinus taeda and Pinus echinata: evidence from morphological and molecular data. Can J For Res 34:2508–2516.

Chunco AJ. 2014. Hybridization in a warmer world. Ecol Evol 4:2019-2031.

Comisión Nacional Forestal. 2009. Inventario Nacional Forestal y de Suelos de México 2004-2009. Comisión Nacional Forestal, Zapopan, Jalisco, México.

Conkle MT, Critchfield WB. 1988. Genetic variation and hybridization of ponderosa pine. In: Ponderosa Pine: the species and its management, Washington State University Cooperative Extension, pp 27-43.

Critchfield WB. 1967. Crossability and relationships of the closed-cone pines. Silvae Genet 16:89-97.

Critchfield WB. 1975. Interspecific hybridization in Pinus: a summary review, pp. 99-105 in: DP Fowler and CY Yeatman (eds.) Symp. on Interspecific and Interprovenance Hybridization in Forest Trees. Proc. 14th Meeting, Canad. Tree Improv. Assoc., Part II.

Debreczy Z, Rácz I. 2011. Conifers around the World. Vols. 1–2. DendroPress Ltd., Budapest. 1089 pp.

Cullingham CI, Cooke JE, Dang S, Davis CS, Cooke BJ, Coltman DW. 2011. Mountain pine beetle host-range expansion threatens the boreal forest. Mol Ecol 20(10):2157-2171.

Delgado P, Salas-Lizana R, Vázquez-Lobo A, Wegier A, Anzidei M, Alvarez-Buylla ER, Piñero D. 2007. Introgressive hybridization in Pinus montezumae Lamb and Pinus pseudostrobus Lindl. (Pinaceae): morphological and molecular (cpSSR) evidence. Int J Plant Sci 168:861-875.

Development Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available: URL http://www.R-project.org/.
Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, et al. 2013. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27-46. DOI: 10.1111/j.1600-0587.2012.07348.x

Dunkey HS. 2001. Pine hybrids: a review of their use performance and genetics. Forest Ecol Manag 148:243-258.

Dvorak WS. 2008. Estamos más cerca de entender la ascendencia de las poblaciones de la "variante patula" en la Sierra Madre del Sur, México? Forest Veracruz 10:59-66.

Dvorak WS, Jordon AP, Hodge GP, Romero JL. 2000. Assessing evolutionary relationships of pines in the Oocarpae and Australes subsections using RAPD markers. New Forests 20:163-192.

Dvorak WS, Kietzka E, Hodge GR, Nel A, Dos Santos GA, Gantz C. 2007. Assessing the potential of Pinus herreræ as a plantation species for the subtropics. Forest Ecol Manag 242:598-605.

Escobar-Flores JG, Lopez-Sanchez CA, Sandoval S, Marquez-Linares MA, Wehenkel C. 2018. Predicting Pinus monophylla forest cover in the Baja California Desert by remote sensing. PeerJ 6:e4603.

Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol 7:574-578.

Farjon A, Styles BT (1997) Pinus (Pinaceae). Flora Neotropica Monograph. New York, NY: The New York Botanical Garden.

Fawcett T 2006. An introduction to ROC analysis. Pattern Recognition Letters 27(8):861-874.

Fogelqvist J, Verkhozina AV, Katyshev AI, Pucholt P, Dixelius C, Rönning-Wästljung AC, Berlin S. 2015. Genetic and morphological evidence for introgression between three species of willows. BMC Evol Biol 15:193, doi: 10.1186/s12862-015-0461-7.

Futuyma DJ. 1998. Evolutionary Biology. Sunderland, Massachussetts, Sinauer Associates Inc.

García-Arévalo A, González-Elizondo MS. 2003. Pináceas de Durango. CONAFOR- Instituto de Ecología, AC. México, D.F. 187 pp.

Garrett PW. 1979. Species hybridization in the genus Pinus. USDA Forest Service, Northeast Forest Experiment Station, Station Research Paper 436.

Gernandt DS, Aguirre Dugua X, Vázquez-Lobo A, Willyard A, Moreno Letelier A, Pérez de la Rosa JA, Piñero D, Liston A. 2018. Multi-locus phylogenetics, lineage sorting, and reticulation in Pinus subsection Australes. Am J Bot 105:711-725.
Gernandt DS, López GG, Ortiz-García S, Liston A. 2005. Phylogeny and classification of *Pinus*. Taxon 4:29-42.

Gernandt DS, Willyard A, Syring JV, Liston A. 2011. The Conifers (Pinophyta), pp. 1-39 in: Plomion C, Bousquet J, Kole C (eds.) Genetics, Genomics and Breeding of Conifers. Science Publishers, CRC Press.

Gómez JM, González-Megías A, Lorite J, Abdelaziz M, Perfectti F. 2015. The silent extinction: climate change and the potential hybridization-mediated extinction of endemic high-mountain plants. Biodivers Conserv 24:1843-1857.

Gompert Z, Buerkle CA. 2016. What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. Evol Appl:909-923, doi:10.1111/eva.12380

González-Elizondo MS, González-Elizondo M, Sørensen PD. 2012a. *Arbutus* bicolor (Ericaceae, Arbuteae), una nueva especie de México. Acta Bot Mex 99:55-72.

González-Elizondo MS, González-Elizondo M, Zamudio S. 2012b. Delimitación taxonómica de *Arbutus mollis* y *A. occidentalis* (Ericaceae). Acta Bot Mex 101:49-81.

González-Elizondo MS, González-Elizondo M, Tena-Flores JA, Ruacho-González L, López-Enríquez IL. 2012c. Vegetación de la Sierra Madre Occidental, México: una síntesis. Acta Bot Mex 100:351-403.

González-Elizondo MS, González-Elizondo M, González LR, Enríquez IL, Rentería FR, Flores JT. 2013. Ecosystems and diversity of the Sierra Madre Occidental. USDA For Serv Proc RMRS-P-67.

Grant PR, Grant BR. 1994. Phenotypic and genetic effects of hybridization in Darwin’s finches. Evolution 48:297-316.

Guo YP, Saukel J, Mittermayr R, Ehrendorfer F. 2005. AFLP analyses demonstrate genetic divergence, hybridization, and multiple polyploidization in the evolution of *Achillea* (Asteraceae-Anthemideae). New Phytol 166:273-290.

Hammer Ø, Harper DAT, Ryan PD. 2001. PAST-palaeontological statistics, ver. 1.89. Palaeontologia electronica 4(9).

Hardy OJ, Charbonnel N, Freville H, Heuertz M. 2003. Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. Genetics 163:1467-1482.

Harrison RG. 1993. Hybrids and hybrid zones: historical perspective, pp 3-12 in: Harrison RG (ed.) Hybrid zones and the evolutionary process. Ithaca, NY: Oxford University Press.
Hernández-León S, Germant DS, Pérez de la Rosa JA, Jardón-Barbolla L. 2013. Phylogenetic relationships and species delimitation in Pinus section Trifoliae inferred from plastid DNA. PLoS One 8(7):e70501.

Hipp AL. 2018. Pharaoh’s Dance: the oak genomic mosaic. PeerJ Preprints 7:e27405v2 DOI:10.7287/peerj.preprints.27405v2.

Hipp AL, Manos PS, Hahn M, Avishai M, Bodénès C, Cavender-Bares J, Crowl A, Deng M, Denk T, Fitz-Gibbon S, Gailing O, González-Elizondo MS, González-Rodríguez A, Grimm GW, Jiang X-L, Kremer A, Lesur I, McVay JD, Plomion C, Rodríguez-Correa H, Schulze E-D, Simeone MC, Sork VL, Valencia-Avalos S. 2019. Genomic landscape of the global oak phylogeny. New Phytol (2019) DOI: 10.1111/nph.16162.

Hongwane P, Mitchell G, Kanzler A, Verryn S, Lopez J, Chirwa P. 2018. Alternative pine hybrids and species to Pinus patula and P. radiata in South Africa and Swaziland. South For 80:301-310 DOI: 10.2989/20702620.2017.1393744.

Jasso-Martínez JM, Machkour-M’Rabet S, Vila R, Rodríguez-Arnaiz R, Castañeda-Sortibrán AN. 2018. Molecular evidence of hybridization in sympatric populations of the Enantia jethys complex (Lepidoptera: Pieridae). PLoS ONE 13(5):e0197116 DOI: 10.1371/journal.pone.0197116.

Kaplan Z, Fehr J. 2007. Molecular evidence for a natural primary triple hybrid in plants revealed from direct sequencing. Ann Bot 99:1213-1222.

Kedharnath S. 1984. Forest tree improvement in India. Proc Ind Acad Sci (Pl Sci) 93: 401-412.

Koerber GR, Anderson PA, Seekamp JV. 2013. Morphology, physiology and AFLP markers validate that green box is a hybrid of Eucalyptus largiflorens and E. gracilis (Myrtaceae). Aust Syst Bot 26:156-166.

Krauss SL. 2000. Accurate gene diversity estimates from amplified fragment length polymorphism (AFLP) markers. Mol Ecol 9:1241-1245.

Lanner RM. 1974. A new pine from Baja California and the hybrid origin of Pinus quadrifolia. Southwest Nat 19:75-95.

LaRue EA, Grimm D, Thum RA. 2013. Laboratory crosses and genetic analysis of natural populations demonstrate sexual viability of invasive hybrid watermilfoils (Myriophyllum spicatum× M. sibiricum). Aquat Bot 109:49-53.
Ledig FT. 1998. Genetic variation in Pinus, pp. 251-280 in: Richardson DM (ed.) Ecology and biogeography of Pinus. Cambridge, U. K.: Cambridge University Press.

Lexer C, Fay MF, Joseph JA, Nica MS, Heinze B. 2005. Barrier to gene flow between two ecologically divergent Populus species, P. alba (white poplar) and P. tremula (European aspen): the role of ecology and life history in gene introgression. Mol Ecol 14:1045-1057.

Linder CR, Risenberg LH. 2004. Reconstructing patterns of reticulate evolution in plants. Am J Bot 91(10):1700-1708 DOI:10.3732/ajb.91.10.1700.

Linder CR, Taha I, Rieseberg LH, Seiler GJ, Snow AA. 1998. Long-term introgression of crop genes into wild sunflower populations. Theor Appl Genet 96:339-347.

Little EL, Righter FL. 1965. Botanical descriptions of forty artificial pine hybrids. US Dept. of Agriculture.

Lopez JJ, Abt RC, Dvorak WS, Hodge GR, Phillips R. 2018. Tree breeding model to assess financial performance of pine hybrids and pure species: deterministic and stochastic approaches for South Africa. New Forests 49:123-142.

López-Upton J, Fiscal VV, Mata JJ, Herrera CR, Hernández JJV. 2001. Hibridación natural entre Pinus oocarpa y P. pringlei. Acta Bot Mex 57:51-66.

Mabaso F, Ham H, Nel A. 2019. Frost tolerance of various Pinus pure species and hybrids. South Forests 81:273-280.

Marie AD, Bernatchez L, Garant D. 2011. Empirical assessment of software efficiency and accuracy to detect introgression under variable stocking scenarios in brook charr (Salvelinus fontinalis). Conservation Genetics 12(5): 1215. doi.org/10.1007/s10592-011-0224-y

Martínez M. 1948. Los pinos mexicanos. 2nd ed Botas, Universidad Nacional Autónoma de México, México, D.F. 361 pp.

McVay JD, Hipp AL, Manos PS. 2017. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. Proceedings of the Royal Society B 284: 20170300.

Menon M, Bagley JC, Friedline CJ, Whipple AV, Schoettle AW, Leal-Saenz A, Wehenkel C, Molina-Freaner F, Flores-Renteria L, Gonzalez-Elizondo MS, Sniezko RA, Cushman SA, Waring KM, Eckert AJ. 2018. The role of hybridization during ecological divergence of southwestern white pine (Pinus strobiformis) and limber pine (P. flexilis). Mol Ecol 27:1245-1260 DOI:10.1111/mec.14505.
Menon M, Landguth E, Leal-Saenz A, Bagley J, Schoettle A, Wehenkel C, Flores-Renteria L, Cushman S, Waring K, Eckert A. 2019. Tracing the footprints of a moving hybrid zone under a demographic history of speciation with gene flow. Evol Appl 2019:1-15 DOI: 10.1111/eva.12795.

Mo J, Xu J, Cao Y, Yang L, Yin T, Hua H, Zhao H, Guo Z, Yang J, Shi J. 2019. *Pinus massoniana* introgression hybrids display differential expression of reproductive genes. Forests 10(3):230 DOI: 10.3390/f10030230.

Mueller UG, Wolfenbarger LL. 1999. AFLP genotyping and fingerprinting. Trends Ecol Evol 14(10):389-394.

Nei M. 1972. Genetic distance between populations. Am Nat 106:283-292.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

Nielsen EE, Bach LA, Kotlicki P. 2006. HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. Mol Ecol Notes 6(4):971-973.

Ortiz-Martínez A, Gernandt D. 2016. Species diversity and plastid DNA haplotype distributions of *Pinus* subsection *Australes* (Pinaceae) in Guerrero and Oaxaca. TIP Rev Esp Cienc Quím Biol 19(2):92-101.

Owens GL, Samuk K. 2019. Adaptive introgression during environmental change can weaken reproductive isolation. bioRxiv 553230 DOI: 10.1101/553230.

Paun O, Schönswetter P. 2012. Amplified Fragment Length Polymorphism (AFLP) - an invaluable fingerprinting technique for genomic, transcriptomic and epigenetic studies. Methods Mol Biol 862:75-87 DOI: 10.1007/978-1-61779-609-8_7.

Peakall R, Smouse PE. 2012. GenAIEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537-2539.

Peñaloza-Ramírez JM, González-Rodríguez A, Mendoza-Cuenca L, Caron H, Kremer A, Oyama K. 2010. Interspecific gene flow in a multispecies oak hybrid zone in the Sierra Tarahumara of Mexico. Ann Bot 105:389-399.

Pérez de la Rosa JA. 1998. Promoción de una variedad de pino serotino mexicano a nivel de especie. Bol Inst Bot Univ Guad (IBUG) 5:127-135.

Pérez de la Rosa JA, Vargas Amado G. 2009. *Pinus luzmariae* Pérez de la Rosa (PINACEAE) y su situación actual en el Bosque La Primavera, Zapopan, Jalisco. II Foro de Investigación y
Conservación del Bosque La Primavera, 5 y 6 de marzo de 2009, Centro Universitario de Ciencias Biológicas y Agropecuarias de la Universidad de Guadalajara, México.

Pérez de la Rosa J, Farjon A. 2013. Pinus luzmariae. The IUCN Red List of Threatened Species 2013: e.T42378A2976301 DOI: 10.2305/IUCN.UK.2013-1.RLTS.T42378A2976301

Perry JP. 1991. The pines of Mexico and Central America. Portland, OR. Timber Press.

Petit RJ, Hampe A. 2006. Some evolutionary consequences of being a tree. Ann Rev Ecol Evol Syst 37:187-214 DOI: 10.1146/annurev.ecolsys.37.091305.110215.

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo Á, Lareu M. 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. Frontiers in Genetics 4:98.

Price RA, Liston A, Strauss SH. 1998. Phylogeny and systematics of Pinus, pp 49-68 in: Richardson DM (ed.), Ecology and Biogeography of Pinus. Cambridge University Press, Cambridge, UK.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Quijada A, Liston A, Robinson W, Alvarez-Buylla E. 1997. The ribosomal ITS region as a marker to detect hybridization in pines. Mol Ecol 6:995-996.

R Development Core Team. 2018 R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org

Rhymer JM, Simberloff D. 1996. Extinction by hybridization and introgression. Ann. Rev Ecol Syst 27:83–109.

Rieseberg LH, Carney SE. 1998. Plant hybridization. New Phytol 140:599-624.

Rieseberg LH, Ellstrand NC, Arnold M. 1993. What can molecular and morphological markers tell us about plant hybridization? Crc Cr Rev Plant Sci 12:213-241.

Sáenz-Romero C, Tapia-Olivares BL. 2003. Pinus oocarpa Isoenzymatic variation along an altitudinal gradient in Michoacán, México. Silvae Genet 52:237-240.

Sáenz-Romero C, Guzmán-Reyna RR, Rehfeldt GE. 2006. Altitudinal genetic variation among Pinus oocarpa populations in Michoacán, Mexico: implications for seed zoning, conservation, tree breeding and global warming. Forest Ecol Manag 229:340-350.
SanCristobal M, Chevalet C, Peleman J, Heuven H, Brugmans B, Van Schriek M, Amigues Y. 2006. Genetic diversity in European pigs utilizing amplified fragment length polymorphism markers. Anim Genet 37:232-238.

Sánchez González A. 2008). Una visión actual de la diversidad y distribución de los pinos de México. Madera y Bosques 14(1):107-120.

Savelainen O, Hedrick P. 1995. Heterozygosity and fitness: no association in Scots Pine. Genetics 140:755-766.

Savelkoul PHM, Aarts HJM, de Haas J, Dijkshoorn L, Duim B, Otsen M, Rademaker JLB, Schouls L, Lenstra JA. 1999. Amplified-fragment length polymorphism analysis: the state of an art. J. Clin. Microbiol. 37: 3083–3091.

Shasany AK, Darokar MP, Dhawan S, Gupta AK, Gupta S, Shukla AK, Khanuja SP. 2005. Use of RAPD and AFLP markers to identify inter-and intraspecific hybrids of Mentha. J Hered 96:542-549.

Simental-Rodríguez SL, Quiñones-Pérez CZ, Moya D, Hernández-Tecles E, López-Sánchez CA, Wehenkel C. 2014. The Relationship between Species Diversity and Genetic Structure in the Rare Picea chihuahuana Tree Species Community, Mexico. PloS ONE 9(11): e111623 DOI: 10.1371/journal.pone.0111623.

Smulders MJM, Beringen R, Volosyanchuk R, vanden Broeck AV, Van der Schoot J, Arens P, Vosman B. 2008. Natural hybridisation between Populus nigra L. and P. x canadensis Moench. Hybrid offspring competes for niches along the Rhine river in the Netherlands. Tree Gen Genom 4:663-675.

Stacy EA, Paritosh B, Johnson MA, Price DK. 2017. Incipient ecological speciation between successional varieties of a dominant tree involves intrinsic postzygotic isolating barriers. Ecol Evol 7:2501-2512 DOI: 10.1002/ece3.2867.

Stewart JF, Liu Y, Tauer CG, Nelson CD. 2010. Microsatellite versus AFLP analyses of pre-management introgression levels in loblolly pine (Pinus taeda L.) and shortleaf pine (P. echinata Mill.). Tree Gen Genom 6(6):853-862 DOI: 10.1007/s11295-010-0296-8.

Strauss SH. 1987. Heterozygosity and developmental stability under inbreeding and crossbreeding in Pinus attenuata. Evolution 41:331-339.
Strobl C, Malley JD, Tutz G. 2009. An Introduction to Recursive Partitioning: Rationale, Application and Characteristics of Classification and Regression Trees, Bagging and Random Forests. Psychol Methods 14:323–348 DOI: 10.1037/a0016973

Styles BT, Stead JW. 1982. Studies of variation in Central American pines. II. Putative hybridization between *Pinus caribaea* var. *hondurensis* and *P. oocarpa*. FAO.

Tatler J, Cassey P, Prowse TA. 2018. High accuracy at low frequency: detailed behavioural classification from accelerometer data. Journal of Experimental Biology 221(23):jeb184085.

Thiers B. 2019. Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. <sweetgum.nybg.org/science/ih/> (accessed Oct 21, 2019).

Tovar-Sánchez E, Oyama K. 2004. Natural hybridization and hybrid zones between *Quercus crassifolia* and *Quercus crassipes* (Fagaceae) in Mexico: morphological and molecular evidence. Am J Bot 91:1352-1363.

Ungerer MC, Baird SJ, Pan J, Rieseberg LH. 1998. Rapid hybrid speciation in wild sunflowers. Proceedings of the National Academy of Sciences 95:11757-11762.

VÄHÄ JP, Primmer CR. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Mol Ecol 15(1):63-72.

Vanden Broeck AV, Storme V, Cottrell JE, Boerjan W, Van Bockstaele E, Quataert P, Van Slycken J. 2004. Gene flow between cultivated poplars and native black poplar (*Populus nigra* L.): a case study along the river Meuse on the Dutch–Belgian border. Forest Ecol Manag 197:307-310.

Vargas-Mendoza CF, Rodríguez-Banderas A, Ibarra-Sánchez CL, Romero-Salas EA, Medina-Jaritz NB, Alcalde-Vázquez R. 2011. Phylogenetic Analysis of Mexican Pine Species Based on Three Loci from Different Genomes (Nuclear, Mitochondrial and Chloroplast), pp. 139-154 in: Agboola JJ (ed.) Relevant Perspectives in Global Environmental Change. InTech. Austria.

Vasilyeva G, Semerikov VL. 2014. Application of amplified fragment length polymorphisms markers to study the hybridization between *Pinus sibirica* and *P. pumila*. Annals of Forest Research 57(2): 175-180 DOI: 10.15287/afr.2014.219.

Vasilyeva G, Goroshkevich S. 2018. Artificial crosses and hybridization frequency in five-needle pines. Dendrobiology 80:123-130 DOI:10.12657/denbio.080.012.
Vekemans X, Hardy OJ. 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. Mol Ecol 13:921-935.

Venables WN, Ripley BD. 1999. Chapter 10: Tree-based methods. In Modern Applied Statistics with S-PLUS. 3rd ed. Eds. J. Chambers, W. Eddy, W. Härdle, S. Sheather, L. Tierney. Statistics and Computing (Springer-Verlag, Press), New York, NY, pp 303-327 DOI: 10.1007/978-1-4757-3121-7_10.

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407-4414.

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Zabeau M. 1999. Two high-density AFLP® linkage maps of Zea mays L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921-935.

Wehenkel C, Simental-Rodríguez LS, Silva-Flores R, Hernández-Díaz C, López-Sánchez CA, Antúnez P. 2015. Discrimination of 59 seed stands of various Mexican pine species based on 43 dendrometric, climatic, edaphic and genetic traits. Forstarchiv 86:194-201.

Wendel JF, McD J, Rettig JH. 1991. Molecular evidence for homoploid reticulate evolution among Australian species of Gossypium. Evolution 45(3):694-711.

Williams CK, Engelhardt A, Cooper T, Mayer Z, Ziem A, Scrucca L, Kuhn MM. 2018. Package ‘caret’.

Wright JW. 1964. Mejoramiento genético de los árboles forestales. FAO. Roma, pp 173-203.

Xu S, Tauer CG, Nelson CD. 2008. Natural hybridization within seed sources of shortleaf pine (Pinus echinata Mill.) and loblolly pine (Pinus taeda L.). Tree Gen Genom 4:849-858.

Zhao W, Meng J, Wang B, Zhang L, Xu Y, Zeng Q-Y, Li Y, Mao J-F, Wang X-R. 2014. Weak crossability barrier but strong juvenile selection supports ecological speciation of the hybrid pine Pinus densata on the Tibetan Plateau. Evolution 68:3120-3133 DOI: 10.1111/evo.12496.
**Table 1** (on next page)

Locations of the stands of *Pinus herrerae* (PH) and *Pinus luzmariae* (PL) under study

Locations of the stands of *Pinus herrerae* (PH) and *Pinus luzmariae* (PL) under study
| Abbreviated stand name | Property | Seed stand | Municipality | Latitude (N) | Longitude (W) | Elevation (m) |
|------------------------|----------|------------|--------------|--------------|--------------|---------------|
| PH-R                   | Comunidad Milpillas | Ranchito | Pueblo Nuevo | 23º 31' | 105º 05' | 2,511 |
| PH-A                   | Comunidad Lajas | Manchon del Abies | Pueblo Nuevo | 23º 11' | 105º 02' | 2,318 |
| PH-V                   | Comunidad Lajas | Ventana | Pueblo Nuevo | 23º 12' | 105º 01' | 2,396 |
| PL-L                   | Comunidad Lajas | Laguna | Pueblo Nuevo | 23º 10' | 105º 07' | 1,960 |
| PL-T                   | Comunidad Lajas | Tacuache | Pueblo Nuevo | 23º 10' | 105º 08' | 2,140 |
Table 2 (on next page)

Cone and needle traits used in the study.

Cone and needle traits of pure *Pinus herrerae* and *Pinus luzmariae* used in the study, min = minimum, max = maximum, SD = Standard deviation.
| Morphological traits                                      | Pinus herrerae | Pinus luzmaria |
|-----------------------------------------------------------|----------------|---------------|
|                                                           | max | mean | min | max | mean | min |
| Cone shape (ovoid (1) vs. widely ovoid (2))               | 1   | 1    | 1   | 2   | 2    | 2   |
| Cone width (cm)                                          | 3.5 | 3.1  | 2.8 | 5.2 | 4.8  | 4.1 |
| Cone scale position (ascendant (1) vs. divergent (3))    | 1   | 1    | 1   | 2   | 2    | 2   |
| Cone scale length (cm)                                   | 1.5 | 1.3  | 1.1 | 1.9 | 1.7  | 1.7 |
| Cone scale width (cm)                                    | 0.8 | 0.7  | 0.6 | 1.1 | 0.9  | 0.8 |
| Apophysis width (mm)                                     | 5   | 4.6  | 4   | 8   | 6.8  | 6   |
| Keel (inconspicuous (0) vs. prominent (1))               | 1   | 0.4  | 0   | 0   | 0    | 0   |
| Leaf sheath length (cm)                                  | 1.3 | 1.1  | 0.9 | 2.0 | 1.8  | 1.5 |
| Leaf sheath diameter (mm)                                | 1.4 | 1.2  | 1.0 | 1.9 | 1.8  | 1.7 |
| Needle number                                            | 3.0 | 3.0  | 3.0 | 3.5 | 3.2  | 3.0 |
| Needle length (cm)                                       | 15.9| 13.4 | 11.3| 28.8| 24.4 | 21.5|
| Needle width (mm)                                        | 1.0 | 0.8  | 0.7 | 1.3 | 1.2  | 1.1 |
| Needle thickness (mm)                                    | 0.3 | 0.2  | 0.1 | 0.6 | 0.5  | 0.4 |
| Stomata rows (dorsal face)                               | 9.0 | 7.1  | 5.3 | 10.0| 9.7  | 9.0 |
| Stomata rows (ventral faces)                             | 3.3 | 3.1  | 3.0 | 6.0 | 5.0  | 4.0 |
Detection of hybrid trees by analysis of 348 AFLP markers

Detection of hybrid trees by analysis of 348 AFLP markers with the software programs STRUCTURE version 2.1 and NewHybrids version 1.1, PH = Pinus herrerai, PL = Pinus luzmareae seed stands, PH-A=Manchon del Abies, PH-R=Ranchito, PH-V=Ventana, PL-L=Laguna and PL-T=Tacuache.
| Seed stand | Sample number | Hybrid number | F$_1$ hybrid number | Hybrid number | F$_1$ hybrid number | Backcrossing number |
|------------|---------------|---------------|---------------------|---------------|---------------------|---------------------|
| PH-A       | 35            | 35            | 2                   | 9             | 0                   | 0                   |
| PH-R       | 33            | 6             | 0                   | 0             | 0                   | 0                   |
| PH-V       | 35            | 14            | 0                   | 2             | 0                   | 1                   |
| PL-L       | 35            | 30            | 1                   | 33            | 0                   | 14                  |
| PL-T       | 33            | 7             | 0                   | 21            | 0                   | 4                   |
| total      | 92            | 3             | 65                  | 0             | 16                  |                     |
Table 4 (on next page)

Accuracy of assignment of *Pinus herrerae*, *Pinus luzmariae* and their hybrids using STRUCTURE 2.1 and NewHybrids 1.1.

Accuracy of assignment of *Pinus herrerae* (PH), *Pinus luzmariae* (PL) and their hybrids using STRUCTURE 2.1 (STR, Pritchard et al., 2000; Falush et al., 2007) and NewHybrids 1.1 (NH, Anderson & Thompson, 2002) using a subset of 11 diagnostic AFLP loci and all 348 polymorphic AFLP loci found in the study. Hybrid classes are as follows: 1\textsuperscript{st} gen – $F_1$, 2\textsuperscript{nd} gen – $F_2$ and $F_1$ backcrosses, and 3\textsuperscript{rd} gen – $F_2$ backcrosses and $F_1$ double backcrosses (Cullingham et al., 2011)
| Class     | 11 AFLP loci |         | 348 AFLP loci |         |
|-----------|--------------|---------|---------------|---------|
|           | STR          | NH      | STR           | NH      |
| 1st gen   | 1            | 1       | 1             | 1       |
| 2nd gen   | 1            | 0.99    | 1             | 0.99    |
| 3rd gen   | 0.59         | 0.49    | 0.50          | 1.00    |
| Hybrid Avg.| 0.86         | 0.83    | 0.83          | 1.00    |
| PH        | 0.13         | 0.88    | 0.80          | 0.88    |
| PL        | 0.13         | 0.88    | 0.78          | 0.95    |
**Table 5** (on next page)

Detection of hybrid trees (*Pinus herreae* x *Pinus luzmariae*) analysis of seven cone and eight needle traits using Random Forest classification.

Detection of hybrid trees (*Pinus herreae* x *Pinus luzmariae*) by analysis of seven cone and eight needle traits using Random Forest classification. PH = *Pinus herreae*, PL = *Pinus luzmariae* seed stands, PH-A=Manchon del Abies, PH-R=Ranchito, PH-V=Ventana, PL-L=Laguna and PL-T=Tacuache.
| Seed stand | Cone traits |   |   | Needle traits |   |   |
|------------|-------------|---|---|--------------|---|---|
|            | Sample number | Hybrid number | Sample number | Hybrid number |
| PH-A       | 35            | 3(7)          | 34            | 0(2)           |
| PH-V       | 34            | 0(10)         | 31            | 2(3)           |
| PL-L       | 31            | 1(3)          | 13            | 0(4)           |
| PL-T       | 31            | 2(9)          | 11            | 2(3)           |
| total      | 131           | 6(29)         | 89            | 4(12)          |

Note: number in brackets = hybrid number at four of seven cone traits detected and at four of eight needle traits detected.
Figure 1

Locations of the *Pinus herrerae* (*Pinus teocote* var. *herrerae*) (yellow circles) and *Pinus luzmariae* stands (red triangles) in the State of Durango, Northwest Mexico.

Locations of the *Pinus herrerae* (*Pinus teocote* var. *herrerae*) (yellow circles) and *Pinus luzmariae* stands (red triangles) in the State of Durango, Northwest Mexico. The *P. herrerae* seed stands were 1) Ranchito (PH-R), Manchon del Abies (PH-A) and Ventana (PH-V). The *P. luzmariae* stands were Laguna (PL-L) and Tacuache (PL-T).
Figure 2

Identification of two populations (K=2) based on 348 AFLP from three *Pinus herrerae* seed stands (PH) (Pop 1 = blue) and two *Pinus luzmariae* seed stands (PL) (Pop 2 = orange) using STRUCTURE.

Identification of two populations (K=2) based on AFLP data from three *Pinus herrerae* seed stands (PH) (Pop 1 = blue) and two *Pinus luzmariae* seed stands (PL) (Pop 2 = red) (171 individuals in total), with Structure, version 2.1 software. PH-A with 35 hybrids, PH-R with six hybrids, PH-V with 14 hybrids, PL-L with 30 hybrids and PL-T with seven hybrids; PH-A=Manchon del Abies, PH-R=Ranchito, PH-V=Ventana, PL-L=Laguna and PL-T=Tacuache
Figure 3

Images from a typical *Pinus luzmariae* (A) and a *Pinus luzmariae* hybrid (B).

Images from a typical *Pinus luzmariae* (Bolaños, Jalisco, 2013) (A) and a *Pinus luzmariae* hybrid (seed stand "Laguna" (PL-L), tree 102, 26 m stem height) (B).
Figure 4

Typical branches, needles and cones of *Pinus herrerae* and *Pinus luzmariae* and variation of *P. luzmariae* cones.

Typical branches, needles and cones of *Pinus herrerae* (A) and *Pinus luzmariae* (B) and variation of *P. luzmariae* cones: typical (C), different hybrid forms (D-G).
Figure 5

Posterior probability ($P$) of being *Pinus herrerae* and *Pinus luzmariae* using Random Forest classification.

Posterior probability ($P$) of being *Pinus herrerae* (PH) and *Pinus luzmariae* (PL) that a tree belongs to a particular class (PH or PL) using a Random Forest classification and A) using seven cone traits, B) using eight needle traits; True Skill Statistic (Allouche, Tsoar & Kadmon, 2006) = +1. If the $P$ of PH (or PL) affiliation of a possible PH (or PL) tree was less than 0.95 (red lines), then, that tree was considered as a putative PH (or PL) hybrid.
Figure 6

Clues of possible hybrid vigour in *Pinus luzmariae*

Clues of possible hybrid vigour in *Pinus luzmariae*: A) Histogram (tree number) of tree heights (m), red line = normal distribution (probability) of all (69) pure *P. luzmariae* trees and putative hybrids under study, green line = logarithm normal distribution of all pure *P. luzmariae* trees and putative hybrids under study, blue bold line = normal distribution of the pure *P. luzmariae* trees under study, B) normal Q-Q-plot of all (69) pure *P. luzmariae* trees and putative hybrids under study.