PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE CIENCIAS EXACTAS Y NATURALES

ESCUELA DE CIENCIAS BIOLÓGICAS

Regional genetic diversity patterns of the palm *Oenocarpus bataua* Mart. as a tool for the conservation of this oleaginous resource in the Neotropics

Tesis previa a la obtención del título de Magíster en Biología de la Conservación

SEBASTIÁN PABLO ESCOBAR VÁSQUEZ

Quito, 2013

Yo, Rommel Montúfar Director de Tesis, CERTIFICO:

Que el señor Sebastián Escobar ha realizado la investigación sobre el tema "Regional genetic diversity patterns of the palm Oenocarpus bataua Mart. as a tool for the conservation of this oleaginous resource in the Neotropics" de acuerdo a las normas y técnicas establecidas. Una vez concluido y revisado el trabajo, conforme con las disposiciones reglamentarias, autorizo la presentación del informe respectivo.

28 de Noviembre del 2013

Rommel Montúfar

DIRECTOR

AGRADECIMIENTOS

En primer lugar deseo agradecer al proyecto PALMS- FP7 (Palms Harvest Impacts in Tropical Forests) y a la Pontificia Universidad Católica del Ecuador (PUCE) por haber financiado esta investigación. Agradezco también al Ministerio del Ambiente del Ecuador por haber concedido los permisos para la realización de este estudio (004-IC-FLO-DPPZ/MA, 06-2012-INVESTIGACIÓN-B-DPMS/MAE, 009-IC-FAU/FLO-DPZCH-MA).

Agradezco especialmente a Rommel Montúfar por su guía durante esta investigación y por las correcciones realizadas a este manuscrito; a Jean-Christophe Pintaud por la coordinación de esta investigación y por el préstamo de tejido vegetal; a Yves Vigouroux, Cédric Mariac, Marie Couderc, Julissa Roncal y a todo el equipo DYNADYV del IRD (Montpellier, Francia) por su ayuda y colaboración durante la fase de laboratorio; a Thomas Couvreur (IRD) por las correcciones adicionales realizadas a este documento.

Agradezco enormemente a Betty Millán de la Universidad Mayor San Marcos (Lima, Perú), a Mónica Moraes de la Universidad Mayor San Andrés (La Paz, Bolivia), y a Jean-Jacques de Granville del IRD (Régina, Guyana Francesa) por su colaboración durante la fase de campo dentro de sus respectivos países. Agradezco en igual medida a Gloria Galeano y a Rodrigo Bernal de la Universidad Nacional de Colombia (Bogotá, Colombia) por el préstamo de tejido vegetal.

Un agradecimiento especial a María José Sanín, Gabriel Rivadeneira y Luis Eduardo López por la ayuda durante la fase de campo, y a Doyle Beaty por las correcciones de idioma efectuadas a este documento.

TABLA DE CONTENIDOS

LISTA DE FIGURAS

LISTA DE TABLAS

El presente trabajo y la literatura citada se encuentran redactados bajo el formato de la revista científica MOLECULAR ECOLOGY.

1. RESUMEN

La diversidad genética dentro de poblaciones ha sido moldeada y estructurada por varios procesos históricos y actuales, tanto a una escala regional como local. Este estudio analiza patrones de diversidad genética regional de *Oenocarpus bataua*, y las fuerzas históricas y actuales que han estructurado su diversidad genética. Esta palmera arborescente se encuentra ampliamente distribuida en la parte septentrional de Sudamérica, sin embargo su divergencia genética intraespecífica permanece conflictual. Dos variedades intraespecíficas y alopátricas, *bataua* (poblaciones occidentales) y *oligocarpa* (poblaciones orientales), han sido descritas. *Oenocarpus bataua* var. *bataua* está ampliamente distribuida dentro de la región del Chocó, la Amazonía occidental y las estribaciones andinas, sin embargo, no se ha reportado ningún tipo de variación morfológica o genética dentro de ésta. Los objetivos de este estudio fueron identificar patrones genéticos intraespecíficos, dentro de la variedad *bataua*, y entre las poblaciones de la variedad *bataua*. Usando siete marcadores SSRs, 644 muestras de cinco países de la región Neotropical (Bolivia, Colombia, Ecuador, Guyana Francesa y Perú) fueron analizadas en este estudio. Usando un análisis de agrupamiento bayesiano, una alta divergencia genética fue identificada entre las poblaciones occidentales y orientales, lo que coincide con la distribución geográfica de *bataua* y *oligocarpa*, respectivamente. Esta divergencia sugiere que ambas variedades podrían ser consideradas como especies distintas con una distribución alopátrica, como fueron anteriormente descritas. Aquí se reporta una localidad peculiar en el norte de Colombia (San Francisco) la cual está relacionada genéticamente con *oligocarpa*, pero se encuentra dentro del rango de distribución de *bataua*. Dentro de *bataua*, cinco agrupamientos genéticos fueron identificados (valle del río Magdalena, Chocó, Amotape-Huancabamba, Napo e Inambari). La divergencia determinada

entre estos agrupamiento está principalmente asociada con el levantamiento de los Andes, lo cual separó la región del Chocó de la Amazonía; y con condiciones climáticas actuales relacionadas con disponibilidad de agua, las cuales han estructurado principalmente a los agrupamientos amazónicos (Napo e Inambari). Altos valores de divergencia genética fueron reportados para los agrupamientos Napo (*He* = 0.83) e Inambari (*He* = 0.79), mientras que los valores más bajos fueron encontrados dentro del agrupamiento del valle del río Magdalena (*He* = 0.54). A una escala local, Intuto (*He* = 0.86) y Jenaro Herrera (*He* = 0.81), en el noreste de Perú, reportaron los valores de diversidad genética más altos. Finalmente, se discute sobre prioridades de conservación para *O. bataua*, inferidas a través de los patrones de diversidad genética reportados en este estudio.

Palabras clave: divergencia genética, diversidad genética, Neotrópico, regiones biogeográficas

2. ABSTRACT

Genetic diversity within populations has been shaped and structured by historical and current processes at a regional and a local scale. This study analyses regional genetic diversity patterns of *Oenocarpus bataua*, and the historical and current forces that have structured its genetic diversity. This arborescent palm is widely distributed in northern South America; however, its intraspecific genetic divergence remains conflictual. Two allopatric varieties, *bataua* (western populations) and *oligocarpa* (eastern populations), have been previously described. *Oenocarpus bataua* var. *bataua* is widely distributed within the Choco region, western Amazonia and Andean foothills, nevertheless, as of yet no intraspecific morphology or molecular variation has been reported. The main goals of this study were to identify genetic patterns at an intraspecific level, within the variety *bataua*, and between variety *bataua* populations. Seven SSRs markers on 644 samples from five countries in the Neotropical region (Bolivia, Colombia, Ecuador, French Guiana and Peru) were analyzed in this study. Using an individual-based Bayesian clustering analysis, high genetic divergence was detected between western and eastern populations which in turn coincides with the geographical distribution of varieties *bataua* and *oligocarpa*, respectively. This divergence suggests that both varieties could be considered distinct species with an allopatric distribution, as formerly described. Here we report a peculiar locality in northern Colombia (San Francisco) genetically related to *oligocarpa*, but within the geographic distribution of *bataua*. Within *bataua*, five genetic clusters were identified (Magdalena river valley, Choco, Amotape-Huancabamba, Napo and Inambari). The major divergence among these genetic clusters is associated with the Andean uplift, which separated the Choco from the Amazonian region; and with current climatic conditions related to water-availability that have principally structured the Amazonian

clusters (Napo and Inambari). High genetic diversity was reported for the Napo (*He* = 0.83) and Inambari (*He* = 0.79) clusters, whereas the lowest diversity was found within the Magdalena river valley cluster (*He* = 0.54). At a local scale, Intuto (*He* = 0.86) and Jenaro Herrera (*He* = 0.81) in northeastern Peru reported the highest levels of genetic diversity. Finally, we discuss conservation priorities for *O. bataua*, inferred from the genetic diversity patterns reported in this study.

Keywords: biogeographic regions, genetic divergence, genetic diversity, Neotropics

3. INTRODUCTION

Our knowledge about species and genetic diversification in the Neotropics is highly influenced by our understanding on how climatic, geological and ecological processes have structured and distributed biodiversity within this region (Cavallotto *et al*. 2011; Eiserhardt *et al*. 2011; Roncal et al. 2013; Turchetto-Zolet *et al*. 2013). Macro-scale processes such as Pleistocene glaciations, lacustrine events, orogenesis and modern climate have shaped present-day diversity patterns of Neotropical species (Wesselingh *et al.* 2006; Lachenaud & Zhang 2008; Eiserhardt *et al*. 2011; Thomas *et al.* 2012). These processes have modified Neotropical landscapes (Storfer *et al.* 2007), promoting vicariant scenarios for genetic diversification (Lachenaud & Zhang 2008; Antonelli & Sanmartín 2011; Turchetto-Zolet *et al.* 2013). Diversity patterns in Neotropical species have also been shaped by micro-scale processes, including topography and habitat heterogeneity (Kerr & Packer 1997; Kristiansen *et al.* 2011), dispersal and other biotic interactions (Kissling *et al.* 2007; Svenning *et al.* 2008), and even by anthropogenic intervention (Morcote-Ríos & Bernal 2001). Since levels of species diversity within communities correlate with levels of genetic diversity within populations in space and time (Vellend & Geber 2005), it is expected that the aforementioned macro and micro-scale processes can also explain intraspecific genetic patterns (Palma-Silva *et al.* 2009).

The palm family (Arecaceae) is one of the oldest monocotyledonous flowering plants (Janssen and Bremer, 2004), and a conspicuous element of tropical and subtropical forests in the Neotropics (Kahn & Granville 1992). In South America, this family is represented by 450 species within 50 genera (Pintaud *et al.* 2008). Palm communities exhibit a high geographical variation in species richness, phylogenetic composition and life forms in this region (Eiserhardt *et al*. 2011). With 36 to 54 different taxa previously reported, western Amazonian

palm communities (Yasuní, Intuto, Jenaro Herrera and Iquitos-Pebas), in particular, harbor the highest amount of diversity (Vormisto *et al.* 2004; Montúfar & Pintaud 2006). Palm species are also locally dominant (Pitman *et al.* 2001; ter Steege et al. 2013); for instance, in only one hectare of primary forest in the Ecuadorian Amazonia, 1255 *Oenocarpus bataua* individuals (at all the stages of development) have been reported (Vormisto *et al*. 2004). Their ecological importance can be addressed by their biotic interactions with pollinators, dispersers and decomposers (Wright & Duber 2001; Galetti *et al.* 2006; Stone 2007; Karubian *et al.* 2010; ThianWoei *et al.* 2012) and by their carbon retention potential (Henson 2005). Additionally, Neotropical palms are considered a major resource, at a socio-economic level, due to the large quantity of uses applied by the local human communities (Brokamp *et al*. 2011). Because of their diversity and abundance, palms are useful models for studying biotic diversification patterns in tropical regions (Bjorholm *et al*. 2005; Svenning *et al*. 2008; Couvreur & Baker 2013).

Oenocarpus bataua is a widely distributed palm species in the tropical and subtropical areas of northern South America and on the island of Trinidad (Henderson 1995). Within its wide geographical distribution, only two intraspecific taxa are recognized: (i) *O. bataua* var. *bataua*, which is located in northwestern South America and found within an altitude gradient from 0 to 1200 m; this variety has been reported within the Choco region, Amazonia basin, Andean forests, and even reported in the Guianas region (Henderson 1995) (ii) *O. bataua* variety *oligocarpa*, distributed in northeastern South America at lower altitudes, up to 580 m (Balick 1986), has been reported in the Guianas region, eastern Venezuela, central Amazonia and on the island of Trinidad (Henderson 1995). The taxonomic separation of these two taxa was based on morphological traits of the flower; such traits include number and localization of

pistillate flowers, length of staminate flowers, number of stamens, and number and diameter of raquillae (Balick 1981; Borgtoft-Pedersen & Balslev 1993). Such morphological differences are supposed to vary within a geographical gradient (Henderson 1995).

Since it was described in 1823 by Martius, *O. bataua* has had a complex history regarding its intraspecific taxonomy (Montúfar & Pintaud 2008). The two current varieties were originally described as distinct species (*O. bataua* and *Jessenia oligocarpa*) (Balick 1986), but were later on considered subspecies (Balick 1986), and finally treated as varieties by Henderson (1995). The morphological stasis of this species, defined by a lack of change in characteristics of gross external anatomy (Bickford *et al*. 2007), was considered an argument valid enough to group them into one single species (Balick 1986; Henderson 1995). However, this morphological-geographical taxonomic model (Henderson 1999) is a simplification of a complex and poorly understood regional genetic structure. For instance, four taxonomic names (*O. seje, J. polycarpa, J. repanda* and *J. weberbaueri*) were described as different *Oenocarpus*/*Jessenia* species; However, these were lately put in synonymy with *O. bataua* var. *bataua* (Berry 1976; Balick 1986; Henderson 1995). Thus, these former taxonomic names could either represent distinct genetic pools, or cryptic diversity hidden within *O. bataua*.

This study explores the genetic variability of *O. bataua* within five countries of South America (Bolivia, Colombia, Ecuador, French Guiana and Peru). The goals of this study were to characterize the genetic diversity of *O. bataua* within three levels: (i) at an intraspecific level, in order to determine genetic differentiation levels between *bataua* and *oligocarpa* types; (ii) within var. *bataua*, to localize genetic clusters within this variety; and (iii) at a population level, so as to identify genetic diversity patterns within var. *bataua*.

4. MATERIALS AND METHODS

4.1. STUDY SPECIES

The palm *O. bataua* is an arborescent, alogamous, monoic species within the tribe Euterpeae. Growing up to around 35 m in height, it reaches the forest canopy, and has a stem diameter of about 20-30 cm wide (Borchsenius *et al*. 1998; Dransfield *et al*. 2008). *O. bataua* is one of the most abundant plant species within the Amazonia with up to 108 individuals per hectare (ter Steege *et al*. 2013), displaying different ecological preferences by growing in many habitats (Montúfar 2008). For instance, in its westernmost distributional limit it grows on well-drained soils (*terra firme* forests) (Borgtoft-Pedersen & Balslev 1993), whereas in the easternmost extreme it grows on poorly drained soils (waterlogged areas, swamps) (Kahn & Granville 1992). The species flowers all year round, but has a peak between February and April in the Amazonian region (Vélez 1992; García 1998), and between December and January in the Colombian Andes (Núñez-Avellaneda & Rojas-Robles 2008). It exhibits thermogenic nocturnal anthesis, and pollination is principally done by arthropods belonging to the orders Coleoptera and Hymenoptera; wind, on the other hand, is poorly involved in pollen dispersion (Núñez-Avellaneda & Rojas-Robles 2008). Although the palms fruits are consumed by birds and mammals, birds are its principal long-distance seed dispersers (Balick 1992; Karubian *et al*. 2010). Furthermore, its fruits and the oil extracted from them are an important alimentary resource for local human populations (Borgtoft-Pedersen & Balslev 1993; Aguilar 2006; Núñez-Avellaneda & Rojas-Robles 2008; Montúfar *et al.* 2010; Brokamp et al. 2011). The oil extracted from the mesocarp of the fruits has a high concentration (> 70%) of oleic fatty acid which is beneficial for human health in nutritional terms (Montúfar *et al.* 2010).

4.2. PLANT MATERIAL

Young leaf and root tissue were collected from a total of 644 *Oenocarpus bataua* individuals at 31 localities during several sampling periods extending from August 2006 to September 2012. After the collection, the samples were stored in plastic bags with silica gel. Five hundred and ninety-one samples were collected from 26 localities in Colombia, Ecuador, Peru and Bolivia, and were identified as var. *bataua*; 53 more samples were collected from six localities in French Guiana and identified as var. *oligocarpa* (Table 1, Fig. 1). Only adult and juvenile individuals were collected. Sampling from closely neighboring plants (distance < 5m) was avoided to prevent temporal and full-siblings sampling bias. Additionally, two individuals of *O. mapora* from Régina (French Guiana) were included as an out-group for phylogenetic analysis.

Sampling of the localities ranged from one (Bahía Málaga) to 57 individuals (Zamora), and was done in a variety of ecosystems, montane (Andes), semi – deciduous (southwestern Ecuador/northern Peru) and tropical (Chocó and Amazonia) forests. Samples were obtained mostly from primary and secondary forests, although some sampling was done in pastures where deforestation has left few individuals standing. All the localities were sampled from primary non-managed forests, although some sampling was done in pastures where deforestation has left few individuals standing. The exception was the Amazonian locality of Chiriap where the seeds of *O. bataua* are dispersed by their habitants within their land.

4.3. DNA EXTRACTION, PRIMER SCREENING AND GENOTYPING

Laboratory procedures were performed in IRD's (Institut de Recherche pour le Développement) facilities in Montpellier, France. DNA extraction was done using the Dneasy plant mini kit from Qiagen following the manufacturer's instructions. After extraction, the DNA concentration of each sample was analyzed with a Thermo Scientific NanoDrop™ ND-100 spectophotometer, and then diluted with ultrapure water to a final concentration of 5 ng/µl.

A total of 35 microsatellite markers (SSR – simple sequence repeats) of, or related to, *O. bataua* were screened (Gaiotto et al. 2001; Lepsch-Cunha et al. 2003; Montúfar et al. 2006). Six of these provided polymorphic SSR data and a strong signal and were therefore selected for further analysis: Ob02, Ob06, Ob08, Ob16, AC5-3#4 and AG5-5#1. One additional intron nuclear microsatellite (AG1) was used for genotyping (Ludeña et al. 2011). In the end, seven microsatellites were used in the genotyping process.

SSR loci were amplified in two separate multiplex polymerase chain reactions (PCR), the first one included four loci (Ob02, Ob06, Ob08, Ob16), while the second one included three (AC5-3#4, AG5-5#1, AG1). Forward primers were fluorescently labeled using VIC, NED, 6-FAM and PET dyes. PCR amplifications were performed in a final volume of 10 µl as follows: 4 µl of DNA template (5 ng/µl), 5 µl of 2X Multiplex Mastermix, and 1 µl of 10X primer mix (containing 3 or 4 pairs of primers at 0.2 µM each primer). Thermal cycling conditions consisted of an initial denaturation step for 15 min at 96 °C followed by 30 cycles of denaturation for 30 s at 94 °C, annealing for 90 s at 56 °C, and elongation for 90 s at 72 °C. A final 10 min extension step at 72 °C was added.

For each PCR product, 2 µl were taken and mixed with 58 µl of water (dilution factor 1/30). Two µl of this PCR diluted product were then mixed with 17 µl of water and 0.15 µl of size marker 500 LIZ (GeneScan™), resulting in a final dilution factor of ~1/300. The final product of each PCR was then analyzed in a 3500 Genetic Analyzer sequencer (Applied Biosystems™) and the resulting chromatograms were examined using *GeneMapper* software v3.7 (Applied Biosystems™).

4.4. STATISTICAL ANALYSIS

The genetic population structure of *Oenocarpus bataua* was analyzed at three different levels: (i) at an intraspecific level (*ISL*), (ii) at an intra var. *bataua* level (*IVBL*), and (iii) at a population level (*PL*). For the first level (*ISL*), the goal was to identify any major genetic divergence (intraspecific categories) within the taxa *O. bataua.* For this level we used five random samples from 20 random localities (with the exception of Guainía and Bahía Málaga where the sampling was smaller) that include the entire geographical distribution of *O. bataua* (including both varieties *bataua* and *oligocarpa*). The final number of samples analyzed was 93. For the second level (*IVBL*), the goal was to identify regional genetic clusters or pools within var. *bataua*. To attain this identification, we worked with all samples identified as var. *bataua* (n = 591) obtained from 26 localities during the fieldwork phase. Within the *PL*, we worked with 19 localities of var. *bataua* (n > 15) and aimed to establish genetic diversity indexes from these populations. For this third level, each locality was treated as a population.

Intraspecific level analysis (*ISL*): A Bayesian clustering method, previously described by Pritchard *et al.* (2000), was used to assign individuals to clusters and determine the genetic

population structure of *O. bataua*. In order to do this, *Structure* v2.3 software was implemented (Pritchard *et al*. 2010). Under this method, the optimal number of genetic clusters (*K*) is usually determined by estimating likelihoods for an increasingly larger *K*, and then choosing the *K* for which the log likelihood reaches a plateau (Pritchard *et al*. 2000). The Δ*K* statistic, developed by Evanno *et al.* (2005), was obtained to infer the optimal number of clusters (*K*).

To determine the optimal number of clusters (*K*), *Structure* was run under three models of ancestry and population intercorrelation: (i) population admixture/correlated allele frequencies, (ii) no admixture/correlated allele frequencies, and (iii) population admixture/independent allele frequencies. Under each model, 5 independent Markov chain Monte Carlo (MCMC) runs were performed using 10^4 burn-in generations followed by 10^5 sampling generations, *K* ranging = 1-10. All other parameters were left at their default values.

To identify areas of genetic discontinuity within the distributional range of *O. bataua*, a spatial Bayesian clustering analysis was performed using *Geneland* software (Guillot *et al.* 2005). Each individual's geo-referenced and genotypic information was used to determine their posterior probability of belonging to a certain cluster. One MCMC run was performed using the resulting *K* value obtained from the *Structure* analysis (following Evanno *et al*.'s method) applying the following parameters: 10^5 iterations, thinning = 1000, allele frequencies correlated, and with uncertainty in the coordinates. For the clusters obtained, inbreeding values (F_{IS}) and F_{ST} values among pairwise clusters were obtained through 10^3 permutations using *Arlequin* 3.1 software (Excoffier et al. 2005).

Intra var. bataua level (*IVBL*) analysis: For this second level of analysis, the number of clusters (*K*) and the identification of areas of genetic discontinuity within var. *bataua* were done as in the former (*ISL*) analysis. Because the San Francisco locality (northwestern Colombia) samples were previously identified in the field as var. *bataua*, we decided to include them in this analysis, despite they were previously grouped along with var. *oligocarpa* samples in the *ISL* analysis. For each cluster, the average number of alleles per locus (*A*), expected heterozygocity (*He*), observed heterozygocity (*Ho*), and inbreeding values (F_{IS}), were obtained with the use of *Arlequin* software (Excoffier *et al.* 2005). F_{ST} values among pairwise clusters were obtained from the same aforementioned software with 10^3 permutations. The number of private alleles per cluster (PA) was obtained with the use of *GenAlEx* 6.5 software (Peakall and Smouse 2012).

Population level (*PL*) analysis: The average number of alleles per locus (*A*), expected heterozygocity (*He*), observed heterozygocity (*Ho*), and inbreeding coefficient (F_{IS}) for each population was estimated with *Arlequin*. *GenAlEx* software was used to identify the number of private alleles per population. To estimate the genetic differentiation between pairwise populations, we implemented Jost's (2008) D_{est} statistical assessment with Meirmans and Hedrick's (2011) statistical correction in *GenAlEx* with 9999 permutations. With the resulting matrix, along with a geographical distance matrix between populations obtained with *ArcMap* 10 software (ESRI 2011), we then performed a Mantel test in *GenAlEx* in order to explore the relationship between genetic structure and geographical space. Phylogenetic relationships between populations were obtained by computing a Neighbour – Joining (NJ) tree with *PopTree2* software (Takezaki *et al.* 2010). The pairwise population distance measure D_A (Nei

et al. 1983) was used as implemented in the software, and bootstrap values were computed with 10^4 replications.

5. RESULTS

5.1. INTRASPECIFIC LEVEL ANALYSIS (*ISL*)

Based on the ΔK statistic (Evanno *et al.* 2005), we determined the best K at $K = 2$ under the three ancestry/allele frequency correlation models used. The $K = 2$ peak made a clear separation between eastern (French Guiana) and western *O. bataua* populations (Choco, Andean and western Amazonian forests) (Fig. 2). This first level of genetic structuring corresponds to the intraspecific variation recognized by taxonomists, which assigns the eastern populations to the var. *oligocarpa*, and the western ones to the var. *bataua*. Under the three ancestry/allele frequency correlation models, a second peak at $K = 4$ appeared. This second peak groups the localities from the var. *oligocarpa* into one cluster, and splits localities of var. *bataua* into three different clusters: (i) San Francisco, as one, (ii) localities from the Choco region, into a second and (iii) the Amazonian localities into a third cluster.

Results produced in *Geneland* (Fig. 2) show a strong assignation of the samples to each corresponding cluster; samples assigned to the *oligocarpa* cluster reported posterior probabilities of membership between 0.9 – 1; meanwhile, samples assigned to the *bataua* cluster reported posterior probabilities of membership between 0.7 – 1. However, an outstanding result was the inclusion of San Francisco to the *oligocarpa* cluster (Fig. 2), which was identified in the field as var. *bataua*. Geographically, var. *oligocarpa* is distributed in eastern Amazonia (Guianas and eastern Venezuela) and has not been reported in northwestern South America where the San Francisco population is located (Magdalena river valley, Colombian western Andean cordillera).

This result increments the potential geographical area of var. *oligocarpa* beyond eastern Venezuela, suggesting the presence of the var. *oligocarpa* even within the Caribbean coast. The F_{ST} value between both clusters (*bataua* y *oligocarpa*) was moderate (0.167) but significantly different from 0; furthermore, F_{IS} values were 0.144 and 0.327 for *bataua* and *oligocarpa* clusters, respectively.

5.2. INTRA VAR. BATAUA LEVEL ANALYSIS (*IVBL*)

In the Bayesian clustering analysis, a peak was identified at $K = 5$ under all the three models of ancestry (Fig. S1). The $K = 5$ peak represents five climatic and geological zones known as the Magdalena river valley (*MRV*), the Choco region (*Choco*), the Amotape-Huancabamba zone (*AHZ*), the Napo region plus northwestern Bolivia (*Napo*) and the *Inambari* region (Inambari) (Fig. 3). Under the second model of ancestry, a second peak appeared at $K = 2$ which represents a grouping of the *Choco* and *AHZ* clusters, the other three clusters were grouped separately. Regarding the third model of ancestry, a second peak appeared at $K = 3$. This K value represents: (i) the *Choco* cluster, (ii) the union of the *Napo*, *AHZ*, and the northern localities of the *Inambari* cluster, and (iii) the *MRV* with the southern localities from both *Napo* and *Inambari* clusters.

The spatial Bayesian clustering analysis showed that the principal genetic barrier was the Andes; however, no visible geographic barriers were identified to explain the separation of the *Napo* and *Inambari* clusters (Fig. 3). Amazonian clusters (*Napo* and *Inambari*) showed the highest genetic diversity (*He* = 0.83 and *He* = 0.79, respectively), whereas the lowest diversity was found at the *MRV* cluster (*He* = 0.54). *Napo* and *Inambari* clusters also reported

the highest *A* value (16.571 and 16.429, respectively) along with the largest amount of samples (n = 186 and n = 155, respectively). The *MRV* cluster, on the contrary, showed the lowest *A* value (*A*=3.429 and n=25) (Table 3). All five clusters showed private alleles, the Amazonian clusters (*Napo* and *Inambari*) being the ones with the most PAs (9 and 10, respectively). All clusters, except the *MRV* cluster which showed a slightly negative F_{IS} value (-0.002), were inbred and displayed positive but moderate F_{IS} values. The highest genetic differentiation (F_{ST} values) was observed in the *MRV* cluster against all the other clusters, especially against the *Choco* cluster (F_{ST} = 0.321), which is supported by *ISL* analysis. *Napo* and *Inambari* clusters had the least amount of genetic differentiation (F_{ST} = 0.043).

5.3. POPULATION – BASED ANALYSIS

The highest genetic diversity was found in the Intuto (*He* = 0.86, A = 11.71) and Jenaro Herrera (*He* = 0.81, *A* = 11.86) populations, which are located close to the town of Iquitos (Peruvian Amazonia), within the *Napo* cluster (Table 4). On the other hand, the lowest genetic diversity was found in northern Colombia (San Francisco), within the *MRV* cluster (*He* = 0.64, *A* = 3.43). Despite its low genetic diversity, San Francisco displayed the most number of private alleles (4). Genetic diversity tends to decrease from western Amazonia to western Ecuadorian Choco populations, and then continues to decrease forward north.

F_{IS} values tended to be low, even negative, with few exceptions (Table 4). High and significant F_{IS} values were coincidently found in populations located near the Andes; However, populations with the highest altitude in southeastern Ecuador showed low F_{IS} values. The highest inbreeding value was observed in southwestern Ecuador, a heavily deforested area.

Pairwise *Dest* values ranged from 0.054 (Chontal vs. Esmeraldas) to 0.985 (San Francisco vs. Villaseca), showing a moderate to strong differentiation between populations. Geographical distances ranged from 53 (Pachicusa vs. Zamora) to 2403 km (San Francisco vs. San Buenaventura). The isolation by distance analysis showed a significant correlation between the genetic and geographical distances of the populations. Fig. 4 shows an increase in genetic divergence between populations conjointly with the increase in geographic distances between them ($p > 0.05$, $r^2 = 0.135$).

The Mantel test showed that the structure observed in the Bayesian clustering and NJ analysis is related with geographical and ecological constraints on gene flow. With few exceptions, the NJ analysis grouped populations in the same pattern as the Bayesian clustering analysis, but with low bootstrap values.

6. DISCUSSION

6.1. INTRASPECIFIC TAXONOMY

This study describes a strong genetic divergence among the varieties *bataua* and *oligocarpa*, as shown in the Bayesian clustering analysis (Fig. 2). The molecular differentiation found between these two varieties is supported by previous studies, which have reported similar patterns using amplified fragments length polymorphism (AFLPs), plastid sequences (*trn*Q-*rps*16, *trn*D-*trn*T, *psb*C-*trn*fM) and fatty acid composition (Montúfar 2007; Montúfar & Pintaud 2008). The molecular evidence presented here draws forth the hypothesis previously reported by the botanists A. Grisebach and H. Wendland, who described the var. *oligocarpa* formerly as *Jessenia*/*Oenocarpus oligocarpa* from the locality type on the island of Trinidad in 1864 (Balick 1986); lately, this taxonomic name has been broadly used to denote eastern Amazonian populations (Suriname, Guianas and Venezuela).

The strong genetic divergence shown in this study suggests that both varieties (*bataua* and *oligocarpa*) could represent cryptic species. Cryptic species are defined as two or more species classified as a single one because they are hard to distinguish morphologically, but can be recognized by molecular data (Bickford *et al*. 2007). Particularly, cryptic diversity has been reported to be more frequent in sympatric scenarios than in allopatric ones (Hebert *et al*. 2004; Stuart *et al*. 2006; Bickford *et al.* 2007). In our case, however, both varieties follow an apparently allopatric model, where each variety is strongly associated with a distinctive geographical and geological region. In this sense, this intraspecific model is hard to fit into a typical cryptic diversity scenario. The morphological variation, particularly of the floral traits, has been poorly studied within the geographical distribution of *O. bataua*. This is mainly due

to the fact that arborescent palms, such as this species, are generally neglected by plant collectors due to their big size (Henderson 1995; Ruokolainen & Vormisto 2000); this generates informational gaps regarding gene flow, gradients on morphological traits and their real distributional ranges. Although *O. bataua* generally follows an apparently allopatric model, the determination of sympatric zones in order to analyze gene flow between the varieties, and the improvement of morphological information through botanical collections will help to elucidate their biological divergence.

The assignment of the San Francisco locality, situated within the Magdalena river valley, to the *oligocarpa* cluster was an outstanding result. Having based our results on Bayesian clustering analysis and genetic differentiation indexes, we determined that this locality showed a peculiar genetic composition which deviates from *bataua*. A possible explanation for this biogeographical pattern lies within three main scenarios: (i) San Francisco could be a hybrid population of *O. bataua* var. *bataua* x *O. bataua* var. *oligocarpa*; nevertheless, it was assigned, and with a high probability (> 0.9), to the *oligocarpa* cluster. Wild hybrids of *O. bataua* var. *bataua* x *O. minor* have been reported near San Francisco (Núñez-Avellaneda 2007), however, its cespitous habit eliminates the possibility of being sampled by mistake. (ii) Botanical studies have reported that the westernmost limit of var. *oligocarpa* is the Miranda Department (north center Venezuela) (Stauffer 1999). However, the geographical range of *oligocarpa* could be wider than what is proposed in the literature, even reaching localities in the Magdalena river valley (northern Colombia). (iii) The San Francisco locality represents a distinctive species within the genus *Oenocarpus*, which is strongly related to *oligocarpa*. This hypothesis, in turn, agrees with the second peak, $K = 4$, observed in the Bayesian clustering analysis (Fig. S1), which defined San Francisco as a unique genetic entity. Therefore, this

locality could also be the result of a bottleneck event of an *oligocarpa* population, where genetic drift has shaped its peculiar genetic footprint. The location of this population coincides with the historical locality type for *Jessenia repanda*, collected and described by botanist H. Engel in 1865 (Balick 1986), and lately put in synonymy with *O. bataua*. This last scenario is particularly more realistic than the two previous.

6.2. BIOGEOGRAPHICAL INFLUENCE ON THE GENETIC STRUCTURE OF *O. bataua* VAR. *bataua*

The five genetic clusters identified in the Bayesian analysis (Fig. 3) correspond to major ecoregions within northwestern South America (Dinerstein *et al*. 1995). The geographical distribution of the clusters correlates, at least partially, with ecological gradients and historical events in this region. Here we discuss the structuring of four of the five clusters; the *MRV* cluster (San Francisco) was previously discussed in the INTRASPECIFIC TAXONOMY section.

Choco region.- The Andean cordillera extends over almost 9000 km along the western coast of South America, and its orogeny has greatly influenced divergence processes in the region (Jørgensen & León-Yánez 1999; Antonelli *et al*. 2010; Chaves *et al*. 2011; Turchetto-Zolet *et al*. 2013). This mountain range is a major physical barrier between the Choco rainforests (trans-Andean region) and the Amazon basin (cis-Andean). It has restricted, totally or at least partially, gene flow between trans/cis-Andean populations; therefore, favoring their genetic identity (Dick *et al.* 2004; Trénel *et al.* 2008). This Andean genetic divergence has been reported for other palm species (*Ceroxylon equinulatunm*, Trénel *et al*. 2008; *Euterpe*

precatoria, Barreiro 2013), rainforest trees (Dick *et al*. 2003; Dick *et al*. 2007; Dick & Heuertz 2008; Hardesty *et al.* 2010; Rymer *et al*. 2013), vertebrates such as frogs, bats and fish (Slade & Moritz 1998; Ditchfield 2000; Perdices *et al.* 2002; Hoffmann & Baker 2003), and arthropods such as butterflies (Brower 1994).

Despite the Andean mountain range being a strong barrier to gene flow, dispersal and pollination processes have occurred and have crossed over this barrier. In this study, genetic differentiation values (F_{ST}) are fairly low between the *Choco* and Amazonian clusters (range: 0.09 – 0.12), suggesting that gene flow can be exchanged between trans/cis-Andean *O. bataua* populations (Montúfar 2007). Studies carried out in other palm species such as *Ceroxylon equinulatum* and *Euterpe precatoria* showed evidence that trans/cis-Andean populations could maintain its genetic connectivity through Andean dispersal corridors located in southern Ecuador/northern Peru region (Trénel *et al.* 2008; Barreiro 2013).

Genetic flux across the Andes, as mentioned above, could have occurred by three non-exclusive scenarios: (i) fluctuation of the Andean tree-line, (ii) historical dispersal and pollination processes through areas of low topography, (iii) current dispersal processes. Pleistocene fluctuations (~2.5 – 0.1 Ma) affecting the altitudinal distribution of Andean forests tree-lines (Van der Hammen 1989) may have reduced the physical isolation between trans/cis-Andean populations in the past. The altitudinal increase of the tree-line in the Andes would have allowed species to ascend altitudinally (Weng *et al*. 2007), reducing the effect of the Andes as a barrier. These altitudinal fluctuations could have favored gene flow events through inter-Andean valleys (Montúfar 2007). Under this scenario, dispersion across the Andes by birds could have played a major role in the movement of alleles (de Queiroz 2005). The second scenario maintains that corridors through Andean valleys, such as the Girón-Paute

deflection in southern Ecuador, the Huancabamba depression in northern Peru or the Caribbean lowlands in northern Colombia, would have allowed continuous gene flow through the Andes (Brumfield & Capparella 1996; Dick *et al.* 2004; Trénel *et al*. 2008). Particularly, pollinators such as bees and bats have been reported to have maintained pollen dispersal through these Andean corridors (Ditchfield 2000; Hoffmann & Baker 2003; Dick *et al*. 2004). Current dispersal processes, however, could also favor gene flow exchange between both sides of the Andes. It is possible that *O. bataua* dispersers maintain some degree of gene flow through these corridors, or even by crossing over the cordillera. For instance, *O. bataua* has long-distance dispersers known as the oildbirds (*Steatornis caripensis)* that could currently maintains some degree of genetic homogeneity between trans/cis-Andean *O. bataua* clusters. It is reported that oilbirds (*Steatornis caripensis)* can travel up to 73.5 km in one day (Snow & Snow 1978; Tannenbaum & Wrege 1978); there is, however, no actual evidence that they transport *O. bataua* seeds across the Andes.

Amotape-Huancabamba zone.- This phytogeographical region, situated between the Río Jubones system (southern Ecuador) and the Río Chamaya system (northern Peru) (Weigend 2002, 2004), is a complex geological zone characterized by high levels of species diversity and endemic species (Jørgensen & Ulloa 1994; Borchsenius 1997; Byg *et al.* 2006; León-Yánez *et al.* 2011). The high biodiversity levels reported in this region are explained by: (i) low species extinction rates due to the absence of Quaternary volcanic activity (Clapperton 1993; Struwe *et al*. 2009), (ii) the overlapping between most northern and southern Andean plant groups (Weigend 2002), and (iii) the mosaic nature of its habitats (Weigend 2002). The high endemicity rates reported are explained by the highly dissected landscape (geodiversity), which may cause small but viable populations to diversify rapidly through drift and/or

adaptive pressures (Weigend 2002). This zone also contains a dispersal corridor for trans/cis-Andean plant groups known as the Huancabamba Depression, an area where the Andean cordillera is interrupted by the Río Chamaya/Río Marañón system located at an altitude of 2145 m (Duellman 1979; Weigend 2002, 2004).

A pattern of genetic structuring similar to that observed in *O. bataua* was also reported for *Theobroma cacao* (Motamayor et al. 2008), where several populations were grouped together into a distinct genetic cluster within this region. Nevertheless, these populations were genetically related with a population in the Choco region, probably due to a genetic exchange through the Huancabamba Depression. Based on molecular and morphological data, a clade within the genus Macrocarpaea (Gentianaceae) was also reported within the location of the Amotape-Huancabamba zone (Struwe *et al*. 2009). This region harbors particular phyogeographical patterns, which are also reflected in the intraspecific genetic diversity of *O. bataua*.

Oenocarpus bataua populations within the *AHZ* cluster are characterized for reaching and passing a 1000 m altitude threshold, which is considered to be the limit between the lowland rainforest and the montane vegetation (Dick *et al*. 2004). Andean *O. bataua* populations, situated at an altitude of over 1000 m, were proposed as a unique ecotype because of their ecological adaptation, which permitted them to grow at such high altitudes (Borgtoft-Pedersen & Balslev 1993). However, the genetic distinctiveness of the *AHZ* cluster is not fully related to altitude, since other Andean populations such as El Chontal (northwestern Ecuador) or Shuaro (central Peru) are genetically related with the *Choco* and *Inambari* clusters, respectively.

Western Amazonia region.- Genetic diversity in western Amazonia was split into two clusters: *Napo* and *Inambari*. The Bayesian clustering analysis recognized these two clusters, although the F_{ST} value between them was less sensitive identifying this pattern (F_{ST} = 0.03). This genetic split is not explained by a modern physical barrier; rather, it could be explained through historical climatic changes such as Pleistocene refuges (Haffer 2008) or by current climate features (Eiserhardt *et al.* 2011).

The Pleistocene refuge theory maintains that the high biodiversity patterns within the Amazonia are explained by allopatric/parapatric speciation processes between isolated populations during the Pleistocene, ~2.5 – 0.1 Ma (Haffer 2008). Areas of the Neotropics with high bird endemicity are hypothesized to developed during the Pleistocene due to vicariance events (Cracraft 1985); two of these endemic areas are known as Napo and Inambari (Cracraft 1985), and their location coincide with two Amazonian genetic clusters for *O. bataua* (*Napo* and *Inambari*). Phylogeographical breaks in birds are known to occur between both regions (Burney & Brumfield 2009), which could have contributed to the genetic differentiation of these regions by dispersal constraints.

Modern climate is one of the principal ecological factors that determine diversity patterns in the Amazonia (Gentry 1988; Kreft *et al*. 2004; Eiserhardt *et al*. 2011). Water-related variables such as annual rainfall and seasonal precipitation are the strongest determinants of plant diversity patterns (Bjorholm *et al*. 2005, 2006; Kreft & Jetz 2007; Eiserhardt *et al*. 2011). Localities within the *Napo* region, such as Iquitos (northeastern Peru), are characterized by high annual precipitation, low seasonal precipitation (Kristiansen *et al*. 2011) and high diversity of palm species (Vormisto *et al.* 2004). Another example is the locality of Yasuní, which harbors diverse palm communities (Vormisto *et al*. 2004), is technically aseasonal

because on any given month it receives at least 100 mm of precipitation, and has a mean annual rainfall of about ~3200 mm (Pitman 2000). Puerto Maldonado, within the *Inambari* cluster, on the other hand, is evidently seasonal, with > 80% of the precipitation falling between October and April and with a mean annual rainfall of about ~2300 mm (Pitman 2000). Modern climate determines plant species diversity patterns (Eiserhardt *et al*. 2011), but also as reported here for *O. bataua*, is a strong determinant of the main patterns of genetic diversity in western Amazonian plants.

The split between the north and south sections of the *Napo* cluster (Fig 3) could be due to a lack of sampling on more eastern localities. Both parts of the *Napo* cluster could be linked by unsampled areas in western Brazil. A similar pattern, also due to a lack of sampling, was reported in the regional genetic structure of *Euterpe precatoria* (Barreiro 2013), where localities from northeastern Colombia were genetically related to localities in southeastern Peru.

6.3. LOCAL DETERMINANTS OF GENETIC DIVERSITY LEVELS

Most genetically diverse populations were identified in western Amazonia, near the Iquitos area (Intuto and Jenaro Herrera). Palm species richness (alpha and gamma diversity) is negatively related to seasonal precipitation (Kristiansen *et al*. 2011); therefore, this variable could also have an influence over genetic diversity levels within western Amazonia. Iquitos is an area characterized by low seasonal precipitation (Kristiansen *et al.* 2011), thus, the high levels of genetic diversity found in this area for *O. bataua* are related to water-availability related determinants, which are the strongest forces shaping palm species diversity patterns in

the Neotropics (Bjorholm *et al.* 2005, 2006; Kreft *et al.* 2006; Eiserhardt *et al*. 2011). Another possible explanation for this genetic diversity pattern is the long-term habitat stability of the western Amazonia (Kristiansen *et al.* 2011), especially during glacial periods when climate changes did not affect precipitation regimes in this area (Mayle *et al.* 2004), nor water availability for palms. The highest levels of genetic diversity in *Theobroma cacao* (Sereno *et al.* 2006; Motamayor *et al.* 2008; Thomas *et al.* 2012) and in *Manihot esculenta* (Fregene et al. 2003) were also found in western Amazonian localities. Apparently, the western Amazonian region tends to harbor the highest levels of genetic diversity within Neotropical trees with a wide geographical distribution, such as *O. bataua* and *T. cacao*.

The location of the highest genetic diversity in *O. bataua* reported in this study coincides with the location of highly diverse palm communities in Neotropical forests (Vormisto *et al*. 2004). The link between the levels of genetic diversity and species diversity has also been reported for other plant species such as *Trillium grandiflorum* and *Plantago lanceolata* (Vellend 2004; Odat *et al*. 2010); however, this is not a rule, as observed in the Alps region (Taberlet et al. 2012). Additionally, the locality of La Pedrera (southeastern Colombia; ~350 km from Iquitos) harbors the greatest amount of diversity within the genus *Oenocarpus*, with six described species (Bernal *et al*. 1991; Galeano & Bernal 2010). We do not have genetic diversity data from southeastern Colombian populations, however the high amount of genetic and species diversity reported in this area suggests that the Napo region and its surroundings are a center of diversification for the genus *Oenocarpus*.

On the other hand, the lowest amount of genetic diversity was reported in northern Colombia and within the Choco region. These populations are peripheral within the geographical distribution of *O. bataua*, contrary to Amazonian populations which are more central. Because

of low rates of gene flow and spatial isolation, peripheral populations have been proposed as areas of lower diversity versus more central areas, (Vucetich & Waite 2003; Eckert *et al.* 2008), an hypothesis supported by our results.

No evident spatial patterns of inbreeding among *O. bataua* populations were reported. Inbreeding is an indicative of the ecological health of populations because it reduces genetic diversity within them (Hall *et al*. 1996; Montúfar *et al*. 2011). The highest rate of inbreeding in this study was reported in southern Ecuador (Villaseca), an area heavily fragmented by deforestation, which represents the edge of the species ecological distribution within semideciduos forests. Deforestation promotes the formation of isolated patches of vegetation where gene flow tends to be low; this human-driven activity increases inbreeding levels within populations (Ramanatha Rao & Hodgkin 2002).

Moderate inbreeding values were also reported for several Andean populations and also for an Amazonian population (Yasuní) of *O. bataua*. Modern-day Andean populations are heavily affected by deforestation and habitat fragmentation which highly exposes them to genetic drift and low rates of gene flow. However, populations at the highest altitudes on the Andean slopes (Zamora and Pachicusa) showed low inbreeding values, suggesting the existence of genetic connectivity mechanisms despite habitat fragmentation. On the contrary, the Yasuní population, located within a primary forest, also reported moderate inbreeding levels; therefore, deforestation is not a unique explanation for the inbreeding rates observed within *O. bataua* populations.

6.4. IMPLICATIONS OF GENETIC DIVERSITY IN CONSERVATION

Intraspecific genetic diversity patterns allow the development of conservation and management actions for wild species (Lemes *et al*. 2003; Eckert *et al*. 2008 Baskauf & Burke 2009). Several criteria can be applied to prioritize the conservation of genetic resources, such as genetic diversity (*He*, *A*, F_{IS}) and distinctive genetic patterns. *In-situ* conservation strategies should be addressed to protect as much genetic variability as possible in order to maintain most of the regional genetic diversity of *O. bataua*. Additionally, national conservation strategies should be focused at local scales, addressing the protection of highly diverse populations such as Intuto, Jenaro Herrera, Pucallpa or Iñapari. These populations with the highest levels of genetic diversity are useful in future breeding and domestication programs, therefore, their conservation should be prioritized (Thomas et al. 2012).

Populations with distinctive ecological or genetic patterns are also addressed as conservation priorities (Thomas et al. 2012). The population of Villaseca is located in the ecological limit of the *O. bataua* palm distribution; therefore, it could display adaptive traits to this habitat. Populations within the *AHZ* cluster, such as Zamora or Rioja, represent a singular genetic pool adapted to high altitudes (> ~1000 m) and are an important oleaginous resource not yet exploited by Andean human communities. Peripheral populations, such as San Francisco (northern Colombia), are potential sites of genetic divergence because of reduced gene flow with more central populations (Lesica & Allendorf 1995). Finally, highly diverse populations of *O. bataua* within the *Choco* cluster, such as El Chontal or Bilsa, should be protected in order to maintain its genetic and ecologic variability as a whole.

Particularly, three genetic clusters were identified within Ecuador (*Choco*, *AHZ* and *Napo*). These clusters should be included in conservation programs such as the National Program of Protected Areas (SNAP), in order to conserve their genetic diversity patterns. *Oenocarpus bataua* develops in several habitats within the *Choco* cluster, such as tropical forests in the province of Esmeraldas, montane forests in Imbabura, and semi-deciduos forest in El Oro. Despite it, its genetic variability is poorly protected by national parks. Mache Chindul and Cotacachi Cayapas National Parks located at north of the Choco region, maintain wild populations of *O. bataua*. However, populations located along the Andean foothills and at southwestern of Ecuador are seriously threatened by deforestation and habitat fragmentation. Populations within the *AHZ* cluster are also threatened by deforestation and by a lack of natural regeneration. However, the genetic variability of *O. bataua* within this region is virtually protected by the Podocarpus National Park, but this still requires to be evaluated. Finally, the *Napo* cluster is better represented within the SNAP, particularly within the Yasuní and Sumaco Napo-Galeras National Parks, and within the Cuyabeno Reserve of Faunistic Protection.

Intraspecific genetic diversity studies give insights about diversification processes, and are an important tool for the production of valuable indicators to prioritize conservation and management strategies of wild populations in the Neotropics. Additionally, genetic diversity patterns could be carefully used as a proxy of species diversity. In this way, these kinds of studies are useful in order to elucidate taxonomic issues within species when morphological traits have not been able to completely address it.

7. LITERATURE CITED

Aguilar Z (2006) Influence of the Huaorani on the conservation of *Oenocarpus bataua*, Arecaceae in Yasuni National Park and Biosphere Reserve, Amazonian Ecuador. *Lyonia: a journal of ecology and application*, 10 (2), 83–90.

Antonelli A, Quijada-Mascareñas A, Crawford AJ, Bates JM, Velazco PM, Wüster W (2010) Molecular studies and phylogeography of Amazonian tetrapods and their relation to geological and climatic models. In: *Amazonia, Landscape and Species Evolution: A Look into the Past* (eds Hoorn & Wesselingh), pp. 386–404. Blackwell Publishing

Antonelli A, Sanmartín I (2011) Why are there so many plant species in the Neotropics? *Taxon*, 60 (2), 403–414.

Balick MJ (1981) *Jessenia bataua* and *Oenocarpus* species: native Amazonian palms as new source of edible oil. In: *New Sources of Fats and Oils* (eds Everett *et al*.), pp. 141–155. American Oil Chemists Society.

Balick MJ (1986) *Systematics and economic botany of the* Oenocarpus-Jessenia *(Palmae) complex*. New York Botanical Garden, Bronx, New York.

Balick MJ (1992) Jessenia *y* Oenocarpus*: palmas aceiteras neotropicales dignas de ser domesticadas.* Organización de la Naciones Unidas para la Agricultura y la Alimentación (FAO), Roma, Italia.

Barreiro JM (2013) *Estructura genética de* Euterpe precatoria *(Mart.) en los Andes Tropicales.* Tesis de Maestría, Pontificia Universidad Católica del Ecuador, Quito, Ecuador.

Baskauf C, Burke J (2009) Population Genetics of *Astragalus bibullatus* (Fabaceae) using AFLPs. *Journal of Heredity*, 100 (4), 424–431.

Bernal RG, Galeano C, Henderson A (1991) Notes on *Oenocarpus* (Palmae) in the Colombian

Amazon. *Brittonia*, 43(3), 154-164.

Berry P (1976) *Estudio bibliográfico y taxonómico preliminar sobre la palma "seje".* Dirección General de los Recursos Hidráulicos CODESUR, Caracas, Venezuela.

Bjorholm S, Svenning J-C, Skov F, Balslev H (2005) Environmental and spatial controls of palm (Arecaceae) species richness across the Americas. *Global Ecology and Biogeography*, 14, 423–429.

Bjorholm S, Svenning J-C, Baker WJ, Skov F, Balslev H (2006) Historical legacies in the geographical diversity patterns of New World palm (Arecaceae) subfamilies. *Botanical Journal of the Linnean Society*, 151, 113–125.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, 22 (3), 148–155.

Borchsenius F (1997) Patterns of plant species endemism in Ecuador. *Biodiversity and Conservation*, 6, 379–399.

Borchsenius F, Balslev H, Pedersen H (1998) *Manual to the Palms of Ecuador*. AAU Reports 37. Aarhus University Press, Aarhus, Denmark.

Borgtoft-Pedersen H, Balslev H (1993) Ungurahua – *Jessenia bataua*. In: *Palmas útiles: especies ecuatorianas para agroforestería y extractivismo* (eds Borgtoft-Pedersen &Balslev), pp. 73–98. Ediciones Abya-Yala, Quito, Ecuador.

Brokamp G, Valderrama M, Mittelbach M, Grandez CA, Barfod AS, Weigend M (2011) Trade in Palm Products in North-Western South America. *The New York Botanical Garden*, 77 (4), 571–606.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the

butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 6491–6495.

Brumfield RT, Capparella AP (1996) Historical diversification on birds in north-western South-America: a molecular perspective on the role of vicariant events. *Evolution*, 50 (4), 1607–1624.

Burney CW, Brumfield RT (2009) Ecology Predicts Levels of Genetic Differentiation in Neotropical Birds. *The American Naturalist*, 174 (3), 358–368.

Byg A, Vormisto J, Balslev H (2006) Using the useful: characteristics of used palms in south-eastern Ecuador. *Environment, Development and Sustainability*, 8, 495–506.

Cavallotto LJ, Violante AR, Hernandez-Molina JF (2011) Geological aspects and evolution of the Patagonian continental margin. *Biological Journal of the Linnean Society*, 103, 346–362.

Chaves J, Weir JT, Smith TB (2011) Diversification in *Adelomyia* hummingbirds follows Andean uplift. *Molecular Ecology*, 20, 4564–4576.

Clapperton C (1993) *Quaternary Geology and Geomorphology of South America*. Elsevier, Amsterdam, Holland.

Couvreur TLP, Baker WJ (2013) Tropical rain forest evolution: palms as a model group. *BioMed Central Biology*, 11, 48.

Cracraft J (1985) Historical Biogeography and Patterns of Differentiation within the South American Avifauna: Areas of Endemism. *Ornithological Monographs*, 36, 49–84.

de Queiroz AD (2005). The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution*, 20 (2), 68-73.

Dick CW, Abdul-Salim K, Bermingham E (2003) Molecular systematics analysis reveals

cryptic Tertiary diversification of a widespread tropical rainforest tree. *American Naturalist*, 162, 691–703.

Dick CW, Roubik DW, Gruber KF, Bermingham E (2004) Long-distance gene flow and cross-Andean dispersal of lowland rainforest bees (Apidae: Euglossini) revealed by comparative mitochondrial DNA phylogeography. *Molecular Ecology*, 13, 3775–3785.

Dick CW, Bermingham E, Lemes MR, Gribel R (2007) Extreme long-distance dispersal of the lowland tropical rainforest tree *Ceiba pentandra* L. (Malvaceae) in Africa and the Neotropics. *Molecular Ecology*, 16 (14), 3039–3049.

Dick CW, Heuertz M (2008) The complex biogeographic history of a widespread tropical tree species. *Evolution*, 62 (11), 2760–2774.

Dinerstein E, Olson DM, Graham DJ, Webster AL, Primm SA, Bookbinder MP, Ledec G (1995) *A conservation assessment of the terrestrial ecoregions of Latin America and the Caribbean*. The World Bank, Washington, D.C.

Ditchfield AD (2000) The comparative phylogeography of Neotropical mammals: patterns of intraspecific mitochondrial DNA variation among bats contrasted to nonvolant small mammals. *Molecular Ecology*, 9, 1307–1318.

Dransfield J, Uhl NW, Asmussen CB, Baker WJ, Harley MM, Lewis CE (2008) *Genera Palmarum*. Royal Botanic Gardens, Kew UK.

Duellman, WE (1979) The herpetofauna of the Andes: Patterns of distribution, origin, differentiation, and present communities (eds Duellman). In: *The South American Herpetofauna: Its Origin, Evolution, and Dispersal.* University of Kansas, Museum of Natural History, Monograph 7, Kansas.

Eckert CG, Samis, KE, Lougheed, SC (2008) Genetic variation across species' geographical

ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, 17, 1170–1188.

Eiserhardt WL, Svenning J-C, Kissling WD, Balslev H (2011) Geographical ecology of the palms (Arecaceae): determinants of diversity and distributions across spatial scales. *Annals of Botany*, 108 (8), 1391–1416.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, 14, 2611–2630.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 2005, 47–50.

Gaiotto F, Brondani R, Grattapaglia D (2005) Microsatellite markers for heart of palm-*Euterpe edulis* and *E. oleracea* Mart. (Arecaceae). *Molecular Ecology*, 1 (1-2), 86–88.

Galeano G, Bernal RG (2010) *Palmas de Colombia*. *Guía de campo*. Editorial Universidad Nacional de Colombia, Bogotá, Colombia.

Galetti M, Donatti CI, Pires AS, Guimarães PR Jordano P (2006) Seed survival and dispersal of an endemic Atlantic forest palm: the combined effects of defaunation and forest fragmentation. *Botanical Journal of the Linnean Society*, 151: 141–149.

García M (1998) Observaciones de polinización en *Jessenia bataua* (Arecaceae). Pontificia Universidad Católica del Ecuador, Quito, Ecuador.

Gentry A (1988) Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri Botanical Garden*, 75, 1-34.

Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. *Molecular Ecology Notes*, 5, 712–715.

Haffer J (2008) Hypotheses to explain the origin of species in Amazonia. *Brazilian Journal of Biology*, 68 (4), 917–947.

Hall P, Walker S, Bawa K (1996) Effect of Forest Fragmentation on Genetic Diversity and Mating System in a Tropical Tree, *Pithecellobium elegans*. *Conservation Biology,* 10 (3), 757–768.

Hardesty BD, Dick CW, Hamrick JL, Degen B, Hubbell SP, Bermingham E (2010) Geographic Influence on Genetic Structure in the Widespread Neotropical Tree *Simarouba amara* (Simaroubaceae). *Tropical Plant Biology*, 3, 28–39.

Hebert PD, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 14812–14817.

Henderson A (1995) *The palms of the Amazon*. Oxford University Press, Oxford.

Henderson A (1999) Species, species concepts, and palm taxonomy in the New World. *Memoirs of the New York Botanical Garden*, 83, 21–28.

Henson IE (2005) An assessment of changes in biomass carbon stocks in tree crops and forests in Malaysia. *Journal of Tropical Forest Science*, 17 (2): 279–296.

Hoffmann FG, Baker RJ (2003) Comparative phylogeography of short-tailed bats (Carollia: Phyllostomidae). *Molecular Ecology*, 12, 3403–3414.

Hoorn C, Wesselingh FP, ter Steege H et al. (2010) Amazonia through time: Andean uplift, climate change, landscape evolution and biodiversity. *Science*, 330, 927–931.

Janssen T, Bremer K. 2004. The age of major monocot groups inferred from 800 + rbcL sequences. *Botanical Journal of the Linnean Society*, 146: 385–398.

Jørgensen PM, Ulloa UC (1994) Seed plants of the high Andes of Ecuador-a checklist. *AAU*

Reports, 34, 1–443.

Jørgensen PM, León-Yánez S (1999) Catalogue of the Vascular Plants of Ecuador. Missouri Botanical Garden, St Louis, Missouri.

Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, 17, 4015–4026.

Kahn F, de Granville JJ (1992) Palms in forest ecosystems of Amazonia. Springer-Verlag, Berlin.

Karubian J, Sork VL, Roorda T, Duraes R, Smith T (2010) Destination-based seed dispersal homogenizes genetic structure of a tropical palm. *Molecular Ecology*, 19, 1745–1753.

Kerr JT, Packer L. 1997. Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. *Nature*, 385: 252–254.

Kissling WD, Rahbek C, Böhning-Gaese K (2007) Food plant diversity as broad-scale determinant of avian frugivore richness. *Proceedings of the Royal Society B: Biological Sciences*, 274 (1611): 799–808.

Kreft H, Köster N, Küper W, Nieder J, Barthlott W (2004) Diversity and biogeography of vascular epiphytes in Western Amazonia, Yasuní, Ecuador. *Journal of Biogeography*, 31, 1463–1476.

Kreft H, Sommer JH, Barthlott W (2006) The significance of geographic range size for spatial diversity patterns in Neotropical palms. *Ecography*, 29: 21–30.

Kreft H, Jetz W (2007) Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences*, 104 (14), 5925–5930.

Kristiansen T, Svenning J-C, Pedersen D, Eiserhardt WL, Grández C, Balslev H (2011) Local and regional palm (Arecaceae) species richness patterns and their cross-scale determinants in the western Amazon. *Journal of Ecology*, 99, 1001–1015.

Lachenaud P, Zhang D (2008) Genetic diversity and population structure in wild stands of cacao trees (*Theobroma cacao* L.) in French Guiana. *Annals of Forest Science*, 65 (3): 1.

Lemes N, Gribel R, Proctor J, Grattapaglia D (2003) Population genetic structure of mahogany (*Swietenia macrophylla* King, Meliaceae) across the Brazilian Amazon, based on variation at microsatellite loci: implications for conservation. *Molecular Ecology*, 12: 2875–2883.

Lepsch-Cunha N, Lund CA, Hamilton MB (2003) Isolation and characterization of nuclear microsatellite loci in the tropical arboreal palm *Oenocarpus bacaba* (Arecaceae). *Molecular Ecology Notes*, 3, 435–437.

León-Yánez S, Valencia R, Pitman N, Endara L, Ulloa Ulloa C, Navarrete H (2011) *Libro Rojo de las Plantas Endémicas del Ecuador, 2da. Edición*. Publicaciones del Herbario QCA, Pontificia Universidad Católica del Ecuador, Quito, Ecuador.

Lesica P, Allendorf FW (1995) When Are Peripheral Populations Valuable for Conservation? *Conservation Biology*, 9 (4), 753–760.

Ludeña B, Chabrillange N, Aberlenc-Bertossi F, Adam H, Tregear J, Pintaud J-C (2011) Phylogenetic utility of the nuclear genes *AGAMOUS 1* and *PHYTOCHROME B* in palms (Arecaceae): an example within Bactridinae. *Annals of Botan*y, 108, 1433–1444.

Mayle FE, Beerling DJ, Gosling WD, Bush MB (2004) Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359 (1443): 499–514.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, 11, 5–18.

Montúfar R (2007) *Structure génétique, biochimique, morphologique et écologique de* Oenocarpus bataua *Mart. (Arecaceae): perspectives pour la valorisation durable d'une ressource forestière néotropicale.* Thèse Doctorat, Supagro, Montpellier.

Montúfar R, Pintaud J-C (2006) Variation in species composition, abundance and microhabitat preferences among western Amazonian *terra firme* palm communities. *Botanical Journal of the Linnean Society*, 151, 127–140.

Montúfar R, Pintaud J-C (2008) Estatus taxonómico de *Oenocarpus bataua* (Euterpeae, Arecaceae) inferido por secuencias del ADN cloroplástico. *Revista Peruana de Biología*, 15, 73–78.

Montúfar R, Mariac C, Pham J, Pintaud J-C (2006) Isolation of 23 polymorphic microsatellite loci in the Neotropical palm *Oenocarpus bataua* Martius (Arecaceae). *Molecular Ecology Notes*, 7 (1), 75–78.

Montúfar R, Laffargue A, Pintaud J-C, Hamon S, Avallone S, Dussert S (2010) *Oenocarpus bataua* Mart. (Arecaceae): Rediscovering a Source of High Oleic Vegetable Oil from Amazonia. *Journal of the American Oil Chemists' Society*, 87 (2), 167–172.

Montúfar R, Anthelme F, Pintaud J-C, Balslev H (2011) Disturbance and Resilience in Tropical American Palm Populations and Communities. *The Botanical Review*, 77 (4), 426–461.

Morcote-Ríos G, Bernal R (2001) Remains of palms (Palmae) at archeological sites in the new world: a review. *The Botanical Review*, 67 (3): 309–350.

Motamayor JC, Lachenaud P, da Silva e Mota JW, Loor R, Kuhn DN, Brown JS, Schnell RJ (2008) Geographic and Genetic Population Differentiation of the Amazonian Chocolate Tree (*Theobroma cacao* L). *PLoS ONE*, 3 (10), e3311.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, 19, 153–170.

Núñez-Avellaneda LA (2007) *Evaluación del aislamiento reproductivo entre tres especies simpátricas* Oenocarpus bataua, Oenocarpus mapora *y* Euterpe precatoria *(Palmae:Arecoideae: Euterpeae) en un bosque de los andes en Colombia*. Tesis de Maestría, Universidad Nacional de Colombia. Bogotá. Colombia.

Núñez-Avellaneda LA, Rojas-Robles R (2008) Biología reproductiva y ecología de la polinización de la palma milpesos *Oenocarpus bataua* en los andes colombianos. *Caldasia*, 30 (1), 101–125.

Odat N, Hellwig SH, Jetschke G, Fischer M (2010) On the relationship between plant species diversity and genetic diversity of *Plantago lanceolata* (Plantaginaceae) within and between grassland communities. *Journal of Plant Ecology*, 3(1), 41–48.

Palma-Silva C, Lexer C, Paggi GM, Barbará T, Bered F, Bodanese-Zanettini MH (2009) Range-wide patterns of nuclear and chloroplast DNA diversity in *Vriesea gigantea* (Bromeliaceae), a neotropical forest species. *Heredity*, 103, 503–512.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics*, 28, 2537–2539.

Perdices A, Bermingham E, Montilla A, Doadrio I (2002) Evolutionary history of the genus *Rhamdia* (Teleostei: Pimelodidae) in Central America. *Molecular Phylogenetics and Evolution*, 25, 172–189.

Pintaud J-C, Galeano G, Balslev H, Bernal R, Borchsenius F, Ferreira E, Granville J-J de, Mejía K, Millán B, Moraes M, Noblick L, Stauffer FW, Kahn F (2008) Las palmeras de América del Sur: diversidad, distribución e historia evolutiva. *Revista Peruana de Biología*, 15 (supl. 1), 7–29.

Pitman NCA (2000) A large-scale inventory of two Amazonian tree communities. Ph.D. Dissertation, Duke University, Durham.

Pitman NCA, Terborgh JW, Silman MR, et al. (2001) Dominance and distribution of tree species in upper Amazonian *terra firme* forests. *Ecology*, 82: 2101–2117.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.

Ramanatha Rao V, Hodgkin T (2002) Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell, Tissue and Organ Culture*, 68, 1–19.

Ruokolainen K, Vormisto J (2000) The most widespread Amazonian palms tend to be tall and habitat generalists. *Basic Applied Ecology*, 1, 97–108.

Roncal J, Kahn F, Millan B, Couvreur TLP, Pintaud J-C (2013) Cenozoic colonization and diversification patterns of tropical American palms: evidence from Astrocaryum (Arecaceae). *Botanical Journal of the Linnean Society*, 171, 120–139.

Rymer PD, Dick CW, Vendramin GG, Buonamici A, Boshier D (2013) Recent phylogeographic structure in a widespread 'weedy' Neotropical tree species, *Cordia alliodora* (Boraginaceae). *Journal of Biogeography*, 40, 693–706.

Sereno ML, Albuquerque PSB, Vencovsky R, Figueira A (2006) Genetic diversity and natural population structure of cacao (*Theobroma cacao* L.) from the Brazilian Amazon evaluated by microsatellite markers. *Conservation Genetics*, 7, 13–24.

Slade RW, Moritz C (1998) Phylogeography of *Bufo marinus* from its natural and introduced ranges. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 265, 769–777.

Snow DW, Snow BK (1978) Palm fruits in the Diet of the Oilbird, *Steatornis caripensis*. *Principes*, 22 (3), 107–109.

Stauffer FW (1999) Datos preliminaries sobre para la actualización de la flora de palmas (Arecaceae) en Venezuela. *Acta Botánica Venezuélica*, 22 (1), 77–107.

Stone AI (2007) Responses of squirrel monkeys to seasonal changes in food availability in an eastern Amazonian forest. *American Journal of Primatology, 69*, 142–157.

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2007) Putting the 'landscape' in landscape genetics. *Heredity*, 98, 128–142.

Struwe L, Haag S, Heiberg E, Grant JR (2009) Andean speciation and vicariance in Neotropical *Macrocarpaea* (Gentianaceae-Helieae). *Annals of the Missouri Botanical Garden*, 96, 450–469.

Stuart BL, Inger RF, Voris HK (2006) High level of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. *Biology Letters*, 2, 470–474.

Svenning J-C, Normand S, Skov F (2008) Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, 31, 316–326.

Taberlet P, Zimmermann NE, Englisch T, Tribsch A, Holderegger R, Alvarez N, Niklfeld H, Coldea G, Mirek Z, Moilanen A (2012) Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters*, 15, 1439–1448.

Takezaki N, Nei M, Tamura K (2010) POPTREE2: Software for Constructing Population Trees from Allele Frequency Data and Computing Other Population Statistics with Windows Interface. *Molecular Biology Evolution*, 27 (4), 747–752.

Tannenbaum B, Wrege P (1978) Ecología del guácharo en Venezuela (*Steatornis caripensis*). *Boletín de la Academía de Ciencias Física, Matemáticas y Naturales*, 115, 73–81.

ter Steege H, Pitman NC, Sabatier D, Baraloto C, Salomão RP, Guevara JE, ..., Silman N

(2013) Hyperdominance in the Amazonian Tree Flora. *Science*, 342 (6156), 1243092.

ThianWoei K, Bong CFJ, King JHP, Leong CTS (2012) Biodiversity of termite (Insecta: Isoptera) in tropical peat land cultivated with oil palms. *Pakistan Journal of Biological Sciences*, 15 (3), 108–120.

Thomas E, van Zonneveld M, Loo J, Hodgkin T, Galluzzi G, van Etten J (2012) Present Spatial Diversity Patterns of *Theobroma cacao* L. in the Neotropics Reflect Genetic Differentiation in Pleistocene Refugia Followed by Human-Influenced Dispersal. *PLos ONE*, 7 (10), e47676.

Turchetto-Zolet AC, Pinheiro F, Salgueiro F, Palma-Silva C (2013) Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, 22, 1193–1213.

Van der Hammen T (1989) History of the montane forests of the northern Andes. *Plant Systematics and Evolution*, 162, 109–114.

Vélez (1992) Estudio fenológico de dieci-nueve frutales silvestres utilizados por las comunidades indígenas de la región de Araracuara-Amazonía Colombiana. *Colombia Amazónica*, 6, 135–186.

Vellend M (2004) Parallel effects of land-use history on species diversity and genetic diversity of forest herbs. *Ecology*, 85, 3043–3055.

Vellend M, Geber MA (2005) Connections between species diversity and genetic diversity. *Ecology Letters*, 8, 767–81.

Vormisto J, Tuomisto H, Oksanen J (2000) Palm distribution patterns in Amazonian rainforests of NE Peru: What is the role of topographic variation? (ed Vormisto) In: *Palms in the rainforests of Peruvian Amazonia: Uses and distribution*. Annales Universitatis Turkuensis Sarja – Ser AII OSA 138.

Vormisto J, Svenning J-C, Hall P, Balslev H (2004) Diversity and dominance in palm (Arecaceae) communities in *terra firme* forests in the western Amazon basin. *Journal of Ecology*, 92, 577–588.

Vucetich JA, Waite TA (2003) Spatial patterns of demography and genetic processes across the species' range: null hypotheses for landscape conservation genetics. *Conservation Genetics*, 4, 639–645.

Weigend M (2002) Observations on the Biogeography of the Amotape-Huancabamba Zone in Northern Peru. *Botanical Review*, 68 (1), 38–54.

Weigend M (2004) Additional observations on the biogeography of the Amotape-Huancabamba zone in Northern Peru: Defining the South-Eastern limits. *Revista Peruana de Biología*, 11 (2), 127–134.

Weng C, Hooghiemstra H, Duivenvoorden JF (2007) Response of pollen diversity to the climate-driven altitudinal shift of vegetation in the Colombian Andes. *Philosophical Transactions of the Royal Society B*, 362, 253–262.

Wesselingh FP, Guerrero J, Räsänen M, Romero Pitmann J, Vonhof H (2006) Landscape evolution and depositional processes in the Miocene Amazonian Pebas lake/wetland system: evidence from exploratory boreholes in northeastern Peru. *Scripta Geologica*, 133, 323–361.

Wright SJ, Duber HC (2001) Poachers and forest fragmentation alter seed dispersal, seed survival, and seedling recruitment in the palm *Attalea butyracea*, with implications for tropical tree diversity. *Biotropica* 33: 583–595.

8. FIGURES

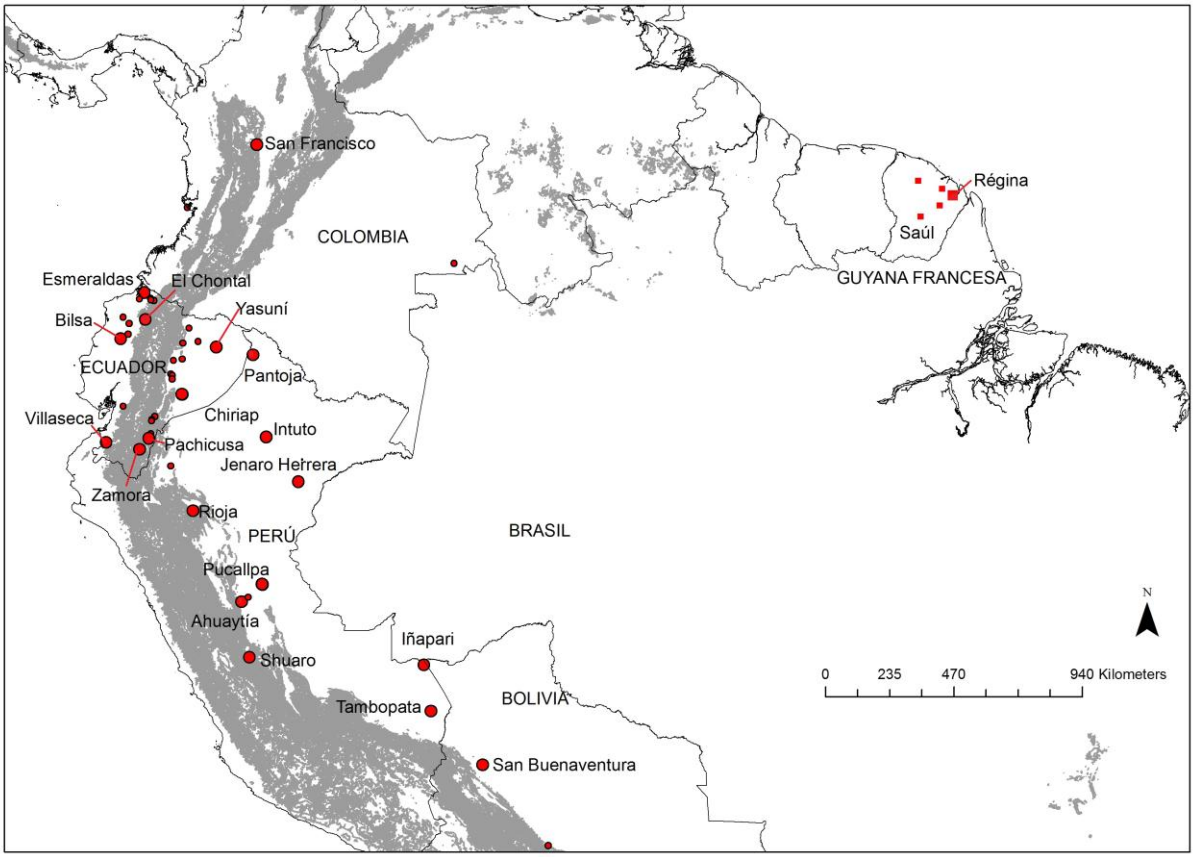

Fig. 1 Locations of the 644 *Oenocarpus bataua* samples. The circles correspond to localities of *Oenocarpus bataua* var. *bataua* while the squares correspond to *Oenocarpus bataua* var. *oligocarpa*. The bigger circumferences and squares represent localities with an n>15. Elevations over 1000 m of altitude are shown in grey.

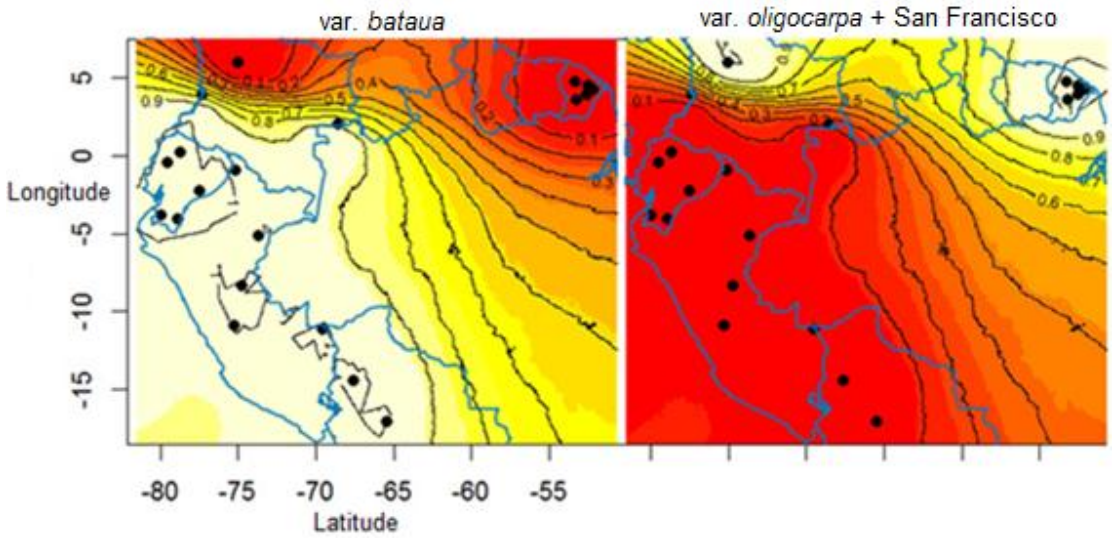

Fig. 2 Genetic clusters identified within *Oenocarpus bataua* (*bataua* and *oligocarpa*) using 93 samples. These were identified with a spatial Bayesian clustering analysis conducted in *Geneland* (Guillot *et al*. 2005) using posterior probabilities to belong to one of K=2 clusters as identified in *Structure* (Pritchard *et al*. 2000) with the statistical analysis developed by Evanno *et al.* (2005). Each point represents a sampled locality while the lines represent the probability of membership to a determined cluster.

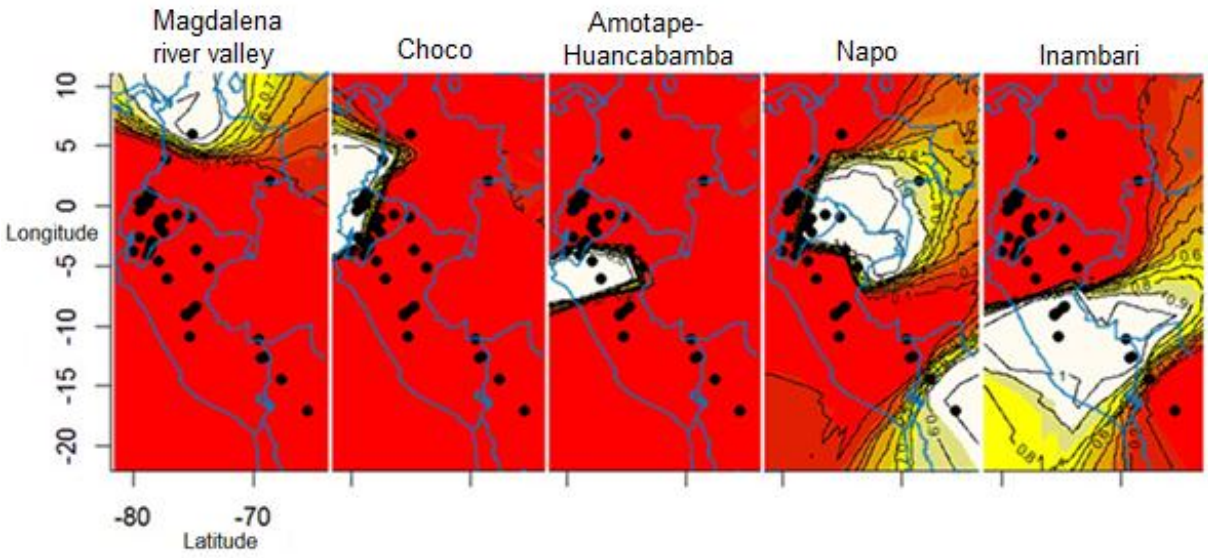

Fig. 3 Genetic clusters identified within *Oenocarpus bataua* var. *bataua* (Magdalena river valley, Choco, Amotape-Huancabamba, Napo and Inambari) using 591 samples. These were identified with a spatial Bayesian clustering analysis conducted in *Geneland* (Guillot *et al.* 2005) using posterior probabilities to belong to one of K=5 clusters as identified in *Structure* (Pritchard *et al.* 2000) with the statistical analysis developed by Evanno *et al*. (2005). Each point represents a sampled locality while the lines represent the probability of membership to a determined cluster.

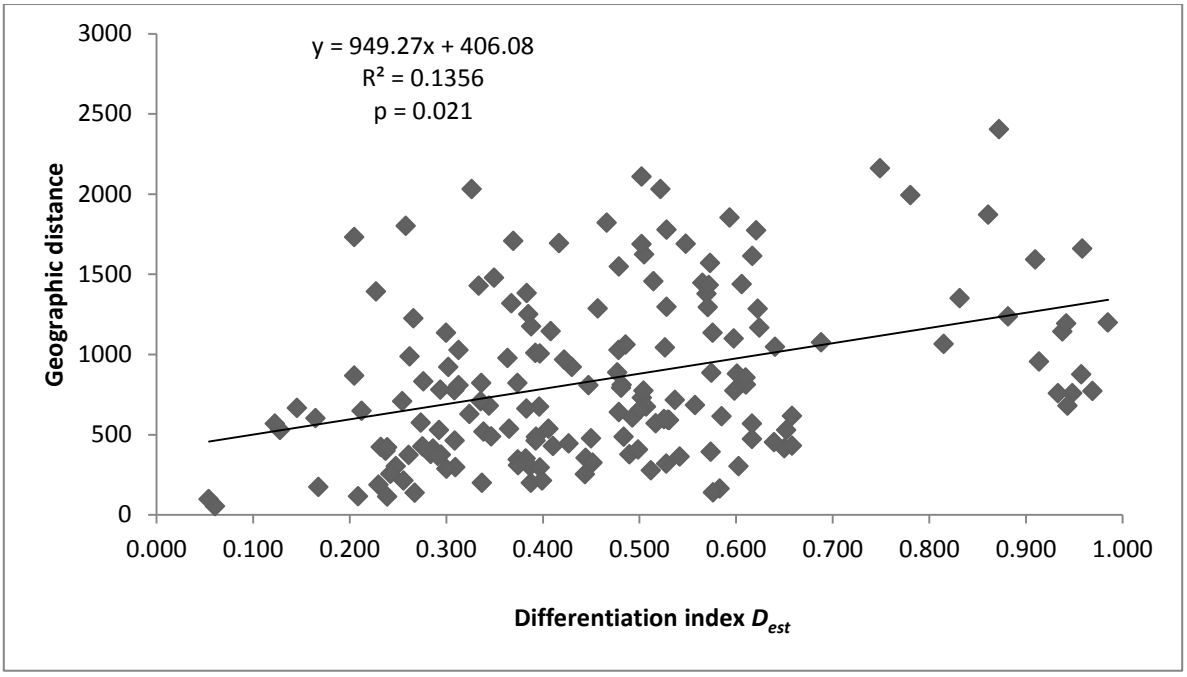

Fig. 4 A Mantel test performed in *GenAlEx* (Peakall & Smouse 2012) with pairwise genetic differentiation values (D_{est}) obtained from *GenAlEx* (Peakall & Smouse 2012) and pairwise geographical distances obtained from *ArcMap* (ESRI 2011). The test was performed with 19 *Oenocarpus bataua* var. *bataua* populations.

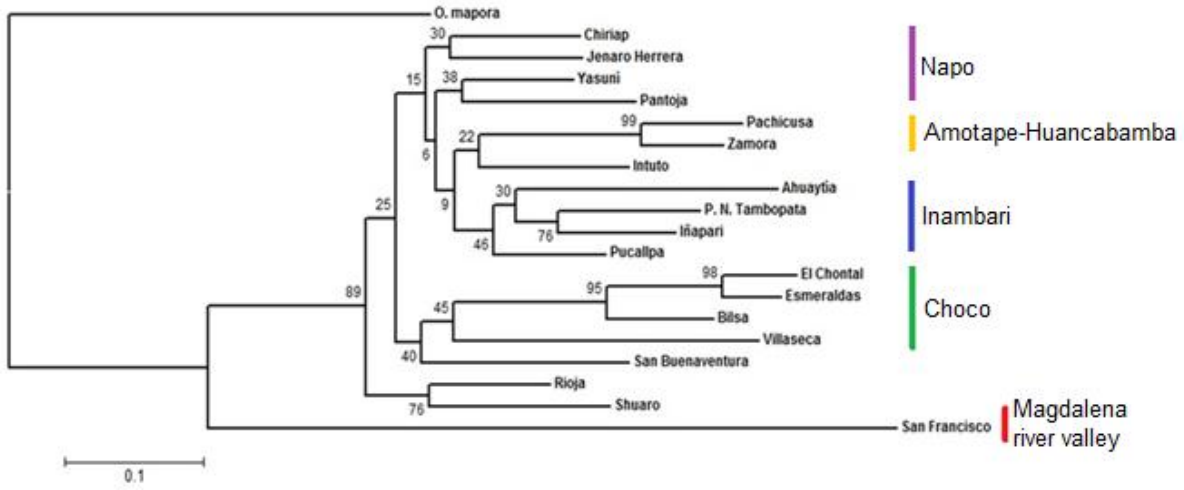

Fig. 5 A Neighbour-Joining analysis was performed in *PopTree2* (Takezaki *et al*. 2010) using 19 *Oenocarpus bataua* var. *bataua* populations. The different colors represent the genetic clusters identified in the Bayesian clustering analysis conducted in Structure (Pritchard et al. 2000).

9. TABLES

Table 1 A general description of the *Oenocarpus bataua* localities sampled within northern South America.

Locality	Variety	n	Latitude	Longitude	Altitude	Forest type
San Francisco	*bataua*	25	5.99336	-75.06756	750	Montane
Bahía Málaga	*bataua*	1	3.92796	-77.34879	7	Chocoan
Guainía	*bataua*	2	2.09109	-68.55888	8	Amazonian
Esmeraldas	*bataua*	26	1.10	-78.71	57	Chocoan
El Chontal	*bataua*	32	0.23419	-78.73266	928	Montane
Bilsa	*bataua*	30	-0.39386	-79.53752	300	Chocoan
Tamarindo	*bataua*	10	-2.6238	-79.45666	512	Chocoan
Villaseca	*bataua*	19	-3.80583	-80.01916	700	Semideciduos
Gualaquiza	*bataua*	6	-3.05	-78.50	1079	Montane
Pachicusa	*bataua*	19	-3.66572	-78.61677	1140	Montane
Zamora	*bataua*	57	-4.03408	-78.92052	1022	Montane
Rioja	*bataua*	18	-6.06169	-77.16831	840	Montane
Tena	*bataua*	6	-1.11	-77.80	500	Amazonian
Yasuní	*bataua*	30	-0.67128	-76.40072	250	Amazonian
Pantoja	*bataua*	26	-0.93277	-75.19194	190	Amazonian
Puyo	*bataua*	7	-1.55	-77.89	980	Montane
Chiriap	*bataua*	30	-2.2331	-77.52628	514	Amazonian
Intuto	*bataua*	32	-3.62273	-74.75322	145	Amazonian
Jenaro Herrera	*bataua*	30	-5.09299	-73.6879	110	Amazonian
Pucallpa	*bataua*	43	-8.35266	-74.78484	165	Amazonian
Ahuaytía	*bataua*	32	-8.96438	-75.62186	349	Amazonian
Shuaro	*bataua*	15	-10.88333	-75.3	800	Montane
Iñapari	*bataua*	30	-11.12727	-69.5596	228	Amazonian
Tambopata	*bataua*	35	-12.58493	-69.04568	194	Amazonian
Carrasco	*bataua*	6	-17.06603	-65.46832	478	Amazonian
San Buenaventura	*bataua*	17	-14.41611	-67.62241	411	Amazonian
Régina	*oligocarpa*	23	4.31666	-52.13333	0	Amazonian
Saül	*oligocarpa*	11	3.61666	-53.2	85	Amazonian
Aratai	*oligocarpa*	8	3.98333	-52.56666	187	Amazonian
Cacao	*oligocarpa*	6	4.52779	-52.49306	22	Amazonian
St. Élie	*oligocarpa*	5	4.79636	-53.27718	188	Amazonian

n=sample size

Table 2 F_{IS} (inbreeding coefficient) and pairwise F_{ST} (fixation index) values obtained from Arlequin (Excoffier *et al.* 2005) for the *bataua* and *oligocarpa* genetic clusters.

			F_{ST}	
Variety	n	F_{IS}	*bataua*	*oligocarpa*
bataua	591	0.144**	0	
oligocarpa	53	0.327**	0.167**	0

n=sample size

Table 3 Diversity and pairwise F_{ST} (fixation index) values obtained from *Arlequin* (Excoffier *et al*. 2005) and *GenAlEx* (Peakall & Smouse 2012) for the five genetic clusters identified in *Structure* (Pritchard *et al*. 2000) within *Oenocarpus bataua* var. *bataua*.

							F_{ST}				
Cluster	n	*A*	*He*	*Ho*	PA	F_{IS}	*MVR*	*CH*	*AHZ*	*N*	*I*
MVR	25	3.43	0.54	0.54	4	-0.002**	0				
CH	118	12	0.75	0.69	3	0.071**	0.32**	0			
AHZ	107	12.57	0.77	0.71	1	0.065**	0.30**	0.12**	0		
N	186	16.57	0.83	0.73	9	0.086**	0.26**	0.09**	0.05**	0	
I	155	16.43	0.79	0.71	10	0.095**	0.27**	0.12**	0.09**	0.04**	0

MRV=Magdalena river valley, *CH*=Choco region, *AHZ*=Amotape-Huancabamba zone, *N* =Napo, *I* = Inambari. n=sample size, *A*=mean number of alleles per locus, *He*=expected heterozygocity, *Ho*=observed heterozygocity, PA=private alleles, F_{IS}=inbreeding coefficient.

Table 4 Diversity values for the 19 *Oenocarpus bataua* var. *bataua* populations obtained from *Arlequin* (Excoffier *et al*. 2005) and *GenAlEx* (Peakall & Smouse 2012).

Locality	Cluster	n	*A*	*He*	*Ho*	PA	F_{IS}
San Francisco	*MRV*	25	3.43	0.54	0.54	4	-0.002
Esmeraldas	*CH*	26	6.43	0.70	0.81	0	-0.158
El Chontal	*CH*	32	6.86	0.70	0.75	0	-0.074
Bilsa	*CH*	31	6.86	0.73	0.84	0	-0.172
Villaseca	*CH*	19	6.6	0.77	0.47	0	0.377**
Pachicusa	*AHZ*	19	7	0.73	0.70	0	0.031
Zamora	*AHZ*	57	8.29	0.71	0.74	0	-0.046
Rioja	*AHZ*	18	8.14	0.76	0.61	0	0.157**
Yasuní	*N*	30	8.43	0.75	0.60	2	0.17**
Pantoja	*N*	26	8.43	0.74	0.65	0	0.038
Chiriap	*N*	30	8.71	0.72	0.88	1	-0.314
Intuto	*N*	32	11.71	0.86	0.82	0	0.034
Jenaro Herrera	*N*	30	11.86	0.81	0.86	2	-0.066
Pucallpa	*I*	43	10.71	0.76	0.75	2	0.007
Ahuaytía	*I*	32	8.14	0.67	0.66	2	0.01
Shuaro	*I*	15	7.14	0.71	0.53	2	0.242**
Iñapari	*I*	30	10.14	0.77	0.82	3	-0.064
Tambopata	*I*	35	9.43	0.72	0.68	0	0.045
San Buenaventura	*N*	17	6	0.75	0.60	1	0.23**

MRV=Magdalena river valley, *CH*=Choco region, *AHZ*=Amotape-Huancabamba zone, *N* =Napo, *I* = Inambari. n=sample size, *A*=mean number of alleles per locus, *He*=expected heterozygocity, *Ho*=observed heterozygocity, PA=private alleles, F_{IS}=inbreeding coefficient.

Table 5 Pairwise genetic differentiation values (D_{est}) between 19 *Oenocarpus bataua* var. *bataua* populations obtained from *GenAlEx* (Peakall & Smouse 2012) are shown beneath the diagonal (all values statistically significance), and the geographic distances obtained from *ArcMap* (ESRI 2011) are shown above the diagonal.

Localities	San Francisco	El Chontal	Esmeraldas	Bilsa	Villaseca	Yasuní	Pantoja	Chiriap	Pachicusa	Zamora	Intuto	Jenaro Herrera	Rioja	Ahuaytía	Pucallpa	Shuaro	Tambopata	Iñapari	San Buenaventura
San Francisco	-	757	679	875	1198	757	771	955	1143	1193	1065	1237	1351	1660	1592	1870	2160	1992	2403
El Chontal	0.948	-	97	113	445	277	415	302	430	472	614	811	716	1074	1046	1285	1774	1613	2030
Esmeraldas	0.943	0.054	-	186	536	324	452	392	529	568	683	885	810	1167	1135	1378	1852	1690	2108
Bilsa	0.957	0.239	0.230	-	354	350	487	302	376	406	640	832	680	1043	1027	1250	1778	1624	2030
Villaseca	0.985	0.427	0.365	0.444	-	518	615	319	163	139	591	730	426	774	792	966	1570	1431	1821
Yasuní	0.933	0.512	0.452	0.383	0.339	-	137	212	413	463	372	574	601	922	868	1135	1547	1382	1801
Pantoja	0.969	0.65	0.64	0.484	0.658	0.268	-	297	486	537	300	488	606	888	822	1099	1456	1286	1707
Chiriap	0.914	0.603	0.574	0.248	0.528	0.256	0.309	-	199	252	344	529	423	774	775	987	1477	1319	1732
Pachicusa	0.938	0.658	0.652	0.49	0.583	0.286	0.393	0.388	-	53	429	570	309	674	673	879	1445	1296	1694
Zamora	0.942	0.617	0.617	0.499	0.576	0.309	0.406	0.444	0.061	-	463	590	295	656	662	856	1438	1295	1688
Intuto	0.815	0.585	0.558	0.479	0.53	0.261	0.387	0.374	0.411	0.393	-	198	375	596	529	807	1175	1008	1428
Jenaro Herrera	0.882	0.61	0.575	0.276	0.503	0.274	0.347	0.128	0.517	0.531	0.337	-	399	476	379	665	976	808	1225
Rioja	0.832	0.537	0.482	0.344	0.276	0.165	0.493	0.232	0.374	0.397	0.295	0.237	-	362	366	568	1144	1005	1392
Ahuaytía	0.959	0.688	0.624	0.527	0.598	0.430	0.477	0.504	0.507	0.501	0.526	0.45	0.542	-	116	213	822	707	1060
Pucallpa	0.91	0.641	0.576	0.479	0.481	0.205	0.336	0.309	0.396	0.383	0.293	0.284	0.291	0.209	-	287	779	647	1027
Shuaro	0.861	0.623	0.57	0.385	0.422	0.3	0.598	0.262	0.601	0.61	0.447	0.146	0.123	0.399	0.3	-	708	627	921
Tambopata	0.749	0.621	0.593	0.528	0.573	0.479	0.515	0.350	0.565	0.606	0.388	0.364	0.408	0.374	0.294	0.336	-	172	255
Iñapari	0.78	0.617	0.548	0.505	0.572	0.383	0.457	0.367	0.528	0.571	0.392	0.313	0.397	0.254	0.213	0.324	0.168	-	421
San Buenaventura	0.872	0.522	0.502	0.326	0.466	0.258	0.37	0.205	0.417	0.502	0.334	0.266	0.227	0.486	0.313	0.302	0.242	0.239	-