



**HAL**  
open science

## Does infraspecific taxonomy match species evolutionary history? A phylogeographic study of *Arundo formosana* (Poaceae)

Laurent Hardion, Regine Verlaque, Maria S. Vorontsova, Isabelle Combroux, Chien-Wen Chen, Tadashi Takamizo, Bruno Vila

### ► To cite this version:

Laurent Hardion, Regine Verlaque, Maria S. Vorontsova, Isabelle Combroux, Chien-Wen Chen, et al.. Does infraspecific taxonomy match species evolutionary history? A phylogeographic study of *Arundo formosana* (Poaceae). *Botanical Journal of the Linnean Society*, 2017, 183 (2), pp.236-249. 10.1093/botlinnean/bow006 . hal-01681572

**HAL Id: hal-01681572**

**<https://hal.science/hal-01681572>**

Submitted on 7 May 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Does infraspecific taxonomy match species evolutionary history? A phylogeographic study of *Arundo formosana* (Poaceae)

LAURENT HARDION<sup>1,2\*</sup>, RÉGINE VERLAQUE<sup>2</sup>, MARIA S. VORONTSOVA<sup>3</sup>, ISABELLE COMBROUX<sup>1</sup>, CHIEN-WEN CHEN<sup>4</sup>, TADASHI TAKAMIZO<sup>5</sup> and BRUNO VILA<sup>2</sup>

<sup>1</sup>Laboratoire Image, Ville, Environnement (LIVE), Université de Strasbourg, CNRS, 67000 Strasbourg, France

<sup>2</sup>Institut Méditerranéen de Biodiversité et d'Ecologie (IMBE), Aix Marseille Université, CNRS, IRD, Avignon Université, 13331 Marseille, France

<sup>3</sup>Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK

<sup>4</sup>Taiwan Forestry Research Institute, Taipei, 10066, Taiwan

<sup>5</sup>National Institute of Livestock and Grassland Science, Ibaraki 305-0901, Japan

Received 8 September 2015; revised 19 February 2016; accepted for publication 11 October 2016

Taxonomic resolution below the species rank often appears challenging, partly due to the recent decline in the popularity of infraspecific groupings in the evolutionary and conservation sciences. We test an integrative approach for reconstructing evolutionary history by reconciling past infraspecific taxonomy with molecular methods, using the Taiwanese endemic *Arundo formosana* (Poaceae). Based on 12 morphometric variables, we provide stronger support for the existence of three morphotypes previously described as varieties and then overlooked, with a clear geographical distribution between western, eastern and northern Taiwan. The phylogeographic analysis of five intergenic regions of plastid DNA supports only the eastern and western lineages and their divergence time according to molecular dating coincides with the orogenesis of the latitudinal central mountain range of Taiwan. AFLP (amplified fragment length polymorphism) nuclear fingerprints also support the east–west divergence in *A. formosana* followed by secondary contacts in the centre of the island, in addition to the monophyly of the northern morphotype nested in the eastern lineage. We suggest an integrative consensus for the taxonomy of *A. formosana* that demonstrates the pertinence of infraspecific taxa in integrative taxonomy and phylogeography below the species level.

ADDITIONAL KEYWORDS: Herbarium material – morphometry – subspecies – systematics – Taiwan – variety.

## INTRODUCTION

Subspecies, variety and form represent the most common taxonomic ranks below the species level (ICBN, McNeil *et al.*, 2012). These infraspecific ranks have traditionally been used as formal delimitations of morphological and geographical variation in species in which taxonomic units were not fully distinct, geographically separate or considered sufficiently different to merit species status. Modern integrative taxonomy aims towards a unified species concept, but has sometimes dismissed infraspecific rank as not relevant to evolutionary history (Sites & Marshall, 2003, 2004; De Queiroz, 2007). The relationship

between lineage history reconstructed by phylogenetic methods, population genetics and groupings below species level remains undefined. The wealth of infraspecific taxa described over the last two centuries has largely not been critically examined using modern methods and exists alongside groups recently defined with the aid of molecular data (i.e. evolutionary significant units, Moritz, 1994). This continued dismissal of infraspecific taxa without critical examination could lead to a loss of diversity knowledge. Taxonomic status also remains influential in biodiversity policy (O'Brien & Mayr, 1991).

Issues concerning the integration of evolutionary biology with taxonomy and conservation have recently been examined in zoology. In avian taxonomy, only 3–36% of subspecies are phylogenetically distinct (Zink, 2004;

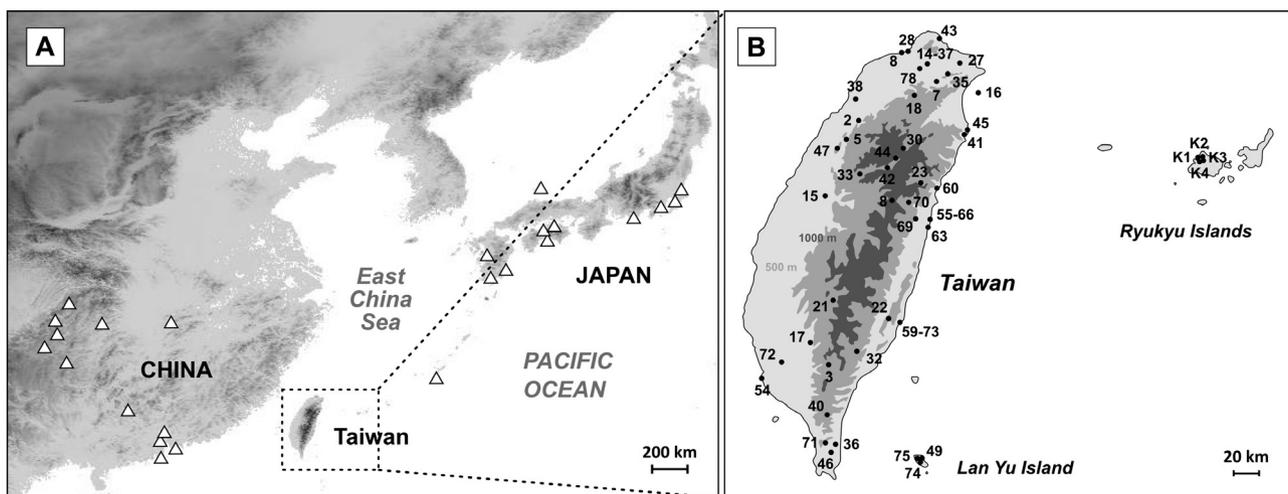
\*Corresponding author. E-mail: [hardion@unistra.fr](mailto:hardion@unistra.fr)

Phillimore & Owens, 2006). This lack of molecular support was also observed among butterflies by Braby, Eastwood & Murray (2012), who called for an integrative concept of subspecies combining phylogenetic support, phenotype distinction and allopatric distribution. In contrast, DNA barcoding of Malaysian butterflies found phylogenetic support for 84% of subspecies (Wilson, Sing & Sofian-Azirun, 2013). The deepest theoretical reflection on infraspecific delimitation in plant taxonomy to date remains the review carried out by Hamilton & Reichard (1992). Based on 494 monographs and revisions published between 1987 and 1990, they found that only 8.2% of the 8043 considered species include infraspecific groups. Among the 661 species with infraspecific delimitations, the authors failed to find any consensus on the scientific reasons cited for the utilization of subspecies (42% of cases) and varietal (52%) ranks, except for the fact that they appear to be preferred by taxonomists from the Old and the New World, respectively. The authors noted the lack of scientific arguments supporting the choice of infraspecific rank, the absence of statistical analyses illustrating patterns of variation within and between ranks and the need for a greater standardization of infraspecific classification. Indeed, numerous modern taxonomic botanists still rely on personal judgement with no formal analysis to define groups of organisms and their appropriate rank (e.g. Vorontsova, Ratovonirina & Randriamboavonjy, 2013).

With 284 species and 58 infraspecific taxa described across > 250 years, *Arundo* L. (Poaceae) represents a suitable model group for testing the evolutionary coherence of infraspecific taxonomy. The majority of taxa originally described in this genus have since been reassigned to other genera (e.g. *Calamagrostis* Adans., *Phragmites* Adans., *Bambusa* Schreb.) and the remaining taxa were

placed in synonymy under just three accepted *Arundo* spp., due to the morphological homogeneity of herbarium specimens (Conert, 1961; Govaerts, 1999). Molecular fingerprinting and morphometric data were used in a recent taxonomic revision of the Mediterranean *A. plinii* Turra s.l. to demonstrate that the radical simplification of taxonomy of *Arundo* has been contrary to the evolutionary history of the group, resulting in the reinstatement of *A. micrantha* Lam. (Hardion et al., 2012a) and the acceptance of another lineage at species level, *A. donaciformis* (Loisel.) Hardion, Verlaque & B.Vila (Hardion et al., 2012a). In addition to the *A. plinii* complex, this genus also includes the subtropical Eurasian *A. donax* L. and the more restricted *A. formosana* Hack. occurring in Taiwan and the Ryukyu Islands (Fig. 1). In his monograph of Arundinae, Conert (1961) recognized three varieties of *A. formosana* based on morphological characters (vars. *formosana*, *gracilis* Hack. and *robusta* Conert), although the infraspecific taxonomy has remained largely ignored in flora treatments (e.g. Huang et al., 2000; Wu, Raven & Hong, 2006b).

The aim of this study is to investigate the evolutionary coherence of infraspecific taxonomy in plant systematics. Subspecies were defined by Mayr (1963) as a 'geographically defined aggregate of local populations which differ taxonomically from other subdivisions of the species', suggesting correspondence to phylogeographic units potentially defined using molecular markers. Do infraspecific units correspond to groups with shared evolutionary and spatial history and morphological similarity or do they represent random variation with no clear morphological, evolutionary or geographical structure? To test these hypotheses, we examine the infraspecific taxonomy of *A. formosana* and its phylogeography in Taiwan, across three complementary datasets: (1) a



**Figure 1.** Geographical distribution of the samples in this study, including (A) Asian samples of *Arundo donax* ( $N = 26$ , white triangles) and (B) the Taiwanese endemic *A. formosana* ( $N = 59$ , black dots). Sampling details are given in Table S1.

morphometric dataset used to test the distinctiveness of *A. formosana* varieties; (2) plastid DNA sequences (including hypervariable sites) used to reconstruct species phylogeny and phylogeography; and (3) AFLP molecular fingerprints generating genome-wide data on genetic clustering, population gene flow and admixture. Clustering patterns from all three datasets are analysed to reassess the taxonomy of *A. formosana* and to bring deeper understanding of lineage history and its relationship with infraspecific taxonomy into the practice of plant systematics. We hope this will encourage the use of modern methods to investigate and reassess other overlooked infraspecific plant taxa.

## MATERIAL AND METHODS

### SAMPLING AND MORPHOMETRIC ANALYSIS

The sampling includes 59 accessions of *A. formosana* from Taiwan and 26 accessions of *A. donax* from eastern Asia, obtained from eight herbaria (BM, E, K, MARS, P, TAIF, US, W; Fig. 1, Table S1). Mediterranean and Middle Eastern populations of *A. donax* were not included in this analysis because of their strongly divergent morphology already documented (under ‘morphotype T1’ in Hardion *et al.*, 2014b). Morphological data were recorded from spikelets collected from the middle part of the panicle, including the following: the number of flowers per spikelet and the length of lower and upper glumes, lemma, palea, lemma awn and hairs. In addition, three characters based on leaf epidermal structures were recorded under a light microscope (Dialux 20, Leitz, Weltzer, Germany): the length of stomatal guard cells, stoma density (per  $10^4 \mu\text{m}^2$ ) and the number of prickles per millimetre of one rib line. To avoid damaging herbarium specimens, leaf epidermis peels were prepared with clear nail polish, following Hilu & Randall (1984), and mounted between a slide and a cover-slide. All ten variables were measured ten times and averaged for each specimen.

The subdivision of groups into discrete taxa has always been a challenging question for groups with complex and continuous variability. Based on phenetic methods, one can expect that the most optimal taxonomic delimitation should maximize intergroup variation (or minimize the within-group variability). This maximum of intergroup variation should be differential between  $k$  and  $k + 1$  groups to avoid the optimal number of groups corresponding to the number of samples. We took this approach using hierarchical clustering on principal components (HCPC; Husson, Josse & Pages, 2010), which can be summarized as three steps: (1) to reduce the morphometric dataset to uncorrelated variables, discovering the main patterns in the data using a principal components analysis (PCA); (2) to construct a classification using an agglomerative

method based on the Ward algorithm and Euclidian distances; and (3) to search for clustering to best explain dataset variation, looking for optimization of the differential intergroup variation between  $k$  and  $k + 1$  clusters along this classification. These analyses were performed in R environment v.2.15 using the HCPC R-function implemented in the *FactorMineR* R-package (R Development Core Team, 2015). They were first performed at an interspecific level between East Asian *A. donax* and *A. formosana* and then at an infraspecific level in *A. formosana*. The coefficient of determination  $R^2$  was calculated for each morphometric variable to illustrate the proportion of the variability explained by taxonomic clustering over the total variability of the data. Finally, each individual of *A. formosana* was assigned to a morphotype using membership probabilities from a discriminant analysis of the morphometric variables (*MASS* R-package; Venables & Ripley, 2002).

### DNA EXTRACTION AND PLASTID DNA SEQUENCING

About 50 mg of leaves from samples collected after 1950 was mechanically crushed after treatment with liquid nitrogen. Total DNA was extracted following Doyle & Doyle (1987) with a modification for herbarium material: incubation of ground material with CTAB isolation buffer (4% CTAB) for 120 min. DNA concentrations were estimated using a Biophotometer (Eppendorf, Hamburg, Germany) and diluted to 50 ng/ $\mu\text{L}$ . Plastid DNA diversity was screened on five intergenic spacers: *trnT-trnL* (Taberlet *et al.*, 1991), *trnCF-rpoB*, *psaA-ORF170*, *rbcL-psaI* and *trnS(GCU)-psbD* (Saltonstall, 2001). Polymerase chain reactions (PCR) followed Hardion *et al.* (2014a): 2.5 mg genomic DNA, 1 $\times$  PCR buffer (10 mM Tris-HCl, 50 mM KCl, 0.001% w/v gelatin), 1.5 mM  $\text{MgCl}_2$ , 2.5 mM each dNTP, 40 pmol each primer, 0.1  $\mu\text{g}/\mu\text{L}$  bovine serum albumin (BSA) and 2.5 units *Taq* polymerase (Q-Biogen, Illkirch, France) in a total volume of 50  $\mu\text{L}$ . The thermal cycling profile was programmed on a PTC-200 Gradient Thermal Cycler (MJ Research, Watertown, MA, USA) as follows: 2 min at 94 °C followed by 35 cycles of 94 °C for 1 min, 56 °C annealing for 1 min, and 72 °C for 2 min, followed by a final extension of 72 °C for 5 min. Purification and sequencing of PCR products were carried out by Eurofins MWG Operon (Ebersberg, Germany).

### PHYLOGEOGRAPHIC ANALYSES

The five plastid DNA regions were manually aligned in MEGA 5.05 (Tamura *et al.*, 2011). Cladistic analyses were performed using maximum-parsimony (MP) to verify phylogenetic congruency of taxa. MP analyses were conducted with PAUP\* v. 4.0b10 (Swofford, 2003), utilizing heuristic searches with TBR branch swapping,

with a time limit of 1 s per search, and 1000 replicates. Bayesian inference (BI) was used to date lineage divergence based on a molecular clock and link these results to the geological formation of Taiwan. Node ages were estimated using BEAST 1.7.4 (Drummond *et al.*, 2006) under a relaxed clock method that assumes a log-normal distribution of rates. Molecular dating is usually calibrated on palaeobotanical records, but these data are of limited availability for grasses due to the uniformity of their pollen and vegetative parts. Consequently, we used dating estimations of divergence between subfamilies of Poaceae calculated from a large dataset in a previous study of molecular dating based on fossil calibration (Christin *et al.*, 2014). The crown node of Arundinoideae, represented by the divergence between *Phragmites* and *Arundo*, was calibrated to 29.0 Ma (confidence interval, 19.6–38.3; Christin, personal communication), following a normal prior distribution. We began the BI analyses with a substitution rate of  $6 \times 10^{-4}$  Ma previously obtained for the phylogeography of another *Arundo* sp. (Hardion *et al.*, 2014a) and estimated it for the four regions. Best-fit substitution models were selected on the Bayesian information criterion (BIC), in jMODELTEST 2.1.1 (Darriba *et al.*, 2012), independently for each region. The tree prior followed the Yule process speciation model to estimate species divergence. Five independent Markov chain Monte Carlo searches were run with  $5 \times 10^7$  generations, sampling every 500 and then discarding 1000 trees for burn-in. Stationarity and convergence of chains were monitored using TRACER 1.5 (Rambaut & Drummond, 2007). The trees were combined with LOGCOMBINER 1.7.4 and a maximum clade credibility tree was summarized in TREEANNOTATOR 17.4 (Drummond *et al.*, 2006). In *A. formosana*, haplotype relationships were inferred using the median-joining network algorithm implemented in Network 4.6 (Bandelt, Forster P, Röhl, 1999). In addition to these cladistic analyses, some insertion/deletion sites (indels) as plastid DNA mini- and microsatellites were included in this analysis for their high ability to resolve phylogeographic patterns (Saltonstall & Lambertini, 2012). Each indel was counted once for each site and they were coded as a fifth state and weighted as 1/10 of a substitution because of their high probabilities of homoplasy (Saltonstall & Lambertini, 2012).

#### AFLP GENOMIC FINGERPRINTS AND ANALYSES

The AFLP procedure followed a slightly modified protocol of Vos *et al.* (1995) described in Hardion *et al.* (2012a): preamplification was performed in 50  $\mu$ L volumes containing 5  $\mu$ L eight-fold diluted ligation product, 10 pmol EcoRI (+A) and MseI (+C) primers, 0.16 mM dNTPs, 0.65 mM MgCl<sub>2</sub> and 1.5 units Taq DNA polymerase (Q-Biogen). The preamplification

thermocycle profile was 94 °C for 2 min, followed by 20 cycles at 94 °C for 45 s, 56 °C for 45 s, 72 °C for 1 min, and 72 °C for 10 min. Two selective PCRs were carried out using EcoRI-AAC/MseI-CAA and EcoRI-AGC/MseI-CTG primer pairs, labelled with 6-FAM fluorescence at the 5' end (Eurofins MWG Operon, Ebersberg, Germany). Selective amplification was performed in 20  $\mu$ L volumes with 5 pmol each primer, 0.65 mM MgCl<sub>2</sub>, 0.5 mM dNTPs, 1 unit of DNA polymerase (Q-Biogen) and 5  $\mu$ L 100 $\times$  diluted preamplification product. The selective amplification thermocycle profile was 94 °C for 2 min, 10 cycles of 94 °C for 30 s, 65 °C for 30 s (touchdown of  $-0.7$  °C per cycle), 72 °C for 1 min, followed by 20 cycles at 94 °C for 30 s, 56 °C for 30 s, 72 °C for 1 min, and 72 °C for 5 min. PCR products were separated and quantified on an ABI 3730xl DNA analyzer (Applied Biosystems, Foster City, CA). AFLP fingerprints were generated from electrophoretogram alignments using the RawGeno R-package (Arrigo *et al.*, 2009). Visualization of the number of fragments per sample in RawGeno was used to select herbarium specimens with the appropriate levels of DNA preservation (i.e. with similar numbers of fragments to the recently dried samples). AFLPs being based on the polymorphism of DNA fragment size, they could be affected by DNA degradation inherent to herbarium conservation. Consequently, we use the number of fragments between 50 and 500 base pairs (bp) per sample as an index of DNA preservation. Genetic markers were automatically selected from the fragments 50–500 bp long, and those exceeding a minimum amplitude of 100 relative fluorescent units were included in the analysis. All unique markers were removed from the dataset. The resulting contingency table was first explored by phenetic methods using the neighbor-joining algorithm with 1000 bootstrap replicates in the R-package ade4 (Dray & Dufour, 2007) and the NeighborNet algorithm using SplitsTree v.4.13 (Huson & Bryant, 2006). The most likely number of genetic clusters was determined by two model-based approaches. The first method utilized discriminant analysis of principal components (DAPC; Jombart, Devillard & Balloux, 2010) using the *find.clusters* R-function implemented in the *adegenet* R-package. This method follows a three-step procedure similar to HCPC described above to determine the optimal number of genetic clusters based on (1) principal components from PCA, (2) sequential *k*-means and (3) model selection minimizing the BIC. The second method uses a Bayesian clustering method implemented in STRUCTURE v.2.3 (Pritchard, Stephens & Donnelly, 2000) under the admixture model, with ten replicates for all values of *K* from 1 to 10, for 1 000 000 steps of which the first 300 000 were discarded as burn-in. The outputs were analysed following Evanno, Regnaut & Goudet (2005) implemented in the R-script Structure-sum-2009 (Ehrich, 2007). STRUCTURE

outputs generated under the admixture model allow the genotypes of each individual to be represented as a mix of ancestral genotypes reconstructed from the molecular dataset, creating a natural way of analysing hybrid zones.

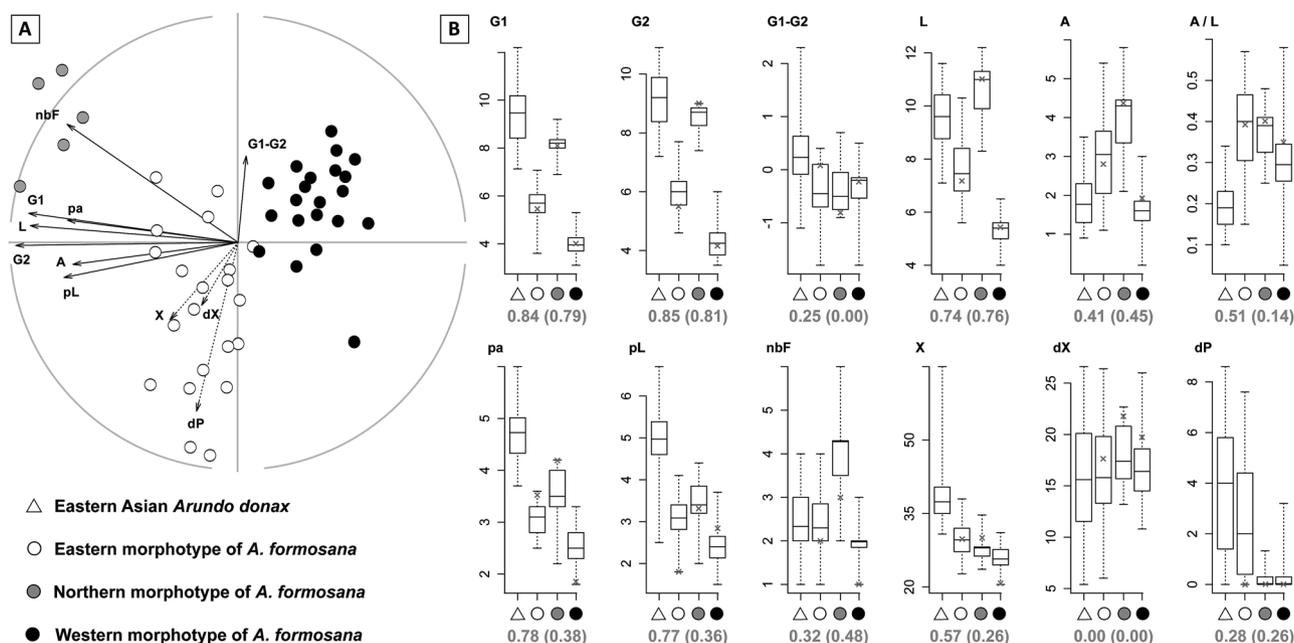
## RESULTS

### MORPHOMETRIC ANALYSES

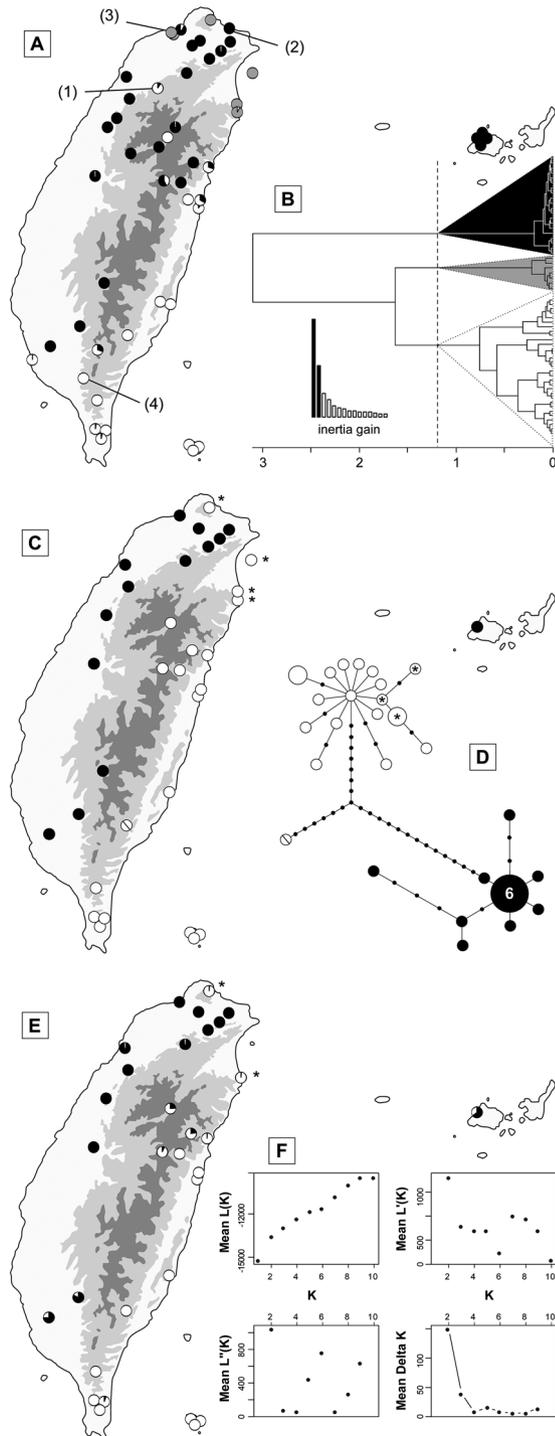
Morphometric analyses demonstrate a clear separation of East Asian *A. donax* from *A. formosana* with two main morphotypes optimized as the best clustering of phenotypic variation by the HCPC analysis (Fig. S1). The three varieties of *A. formosana* are correctly assigned to the cluster of *A. formosana*. This distinction is mainly based on the longer spikelet pieces (i.e. glumes, lemma and palea) of *A. donax* compared to *A. formosana*. East Asian *A. donax* also has more equal glumes compared to the subequal glumes of *A. formosana* (the upper glume longer than the lower).

*A. donax* also has a lemma awn which is shorter in proportion to the lemma length (A/L, Fig. 2).

The best HCPC clustering of phenotypic variation in *A. formosana* divided the accessions into three morphotypes located mainly in eastern, western and northern parts of the island (Fig. 3A, B). Distinct by its longer spikelet, the northern morphotype (including the type of var. *robusta*) is more related to the eastern morphotype (including the type of var. *formosana*) than the western morphotype (including the type of var. *gracilis*) possessing the smallest spikelets. Assignment probabilities from the discriminant analysis applied directly to morphometric variables describe four individuals with equivalent assignment probabilities to the eastern and western morphotypes, three of them being located in the central-east of Taiwan (Fig. 3A). The geographically isolated samples from Ryukyu Islands (K1, K2, K3 and K4) correspond to the western morphotype. The type specimen of *A. parviflora* Ohwi, the only species synonym of *A. formosana* described from southern Taiwan, is related to the eastern morphotype.



**Figure 2.** (A) First plan of the discriminant analysis between *Arundo formosana* morphotypes defined by hierarchical clustering of principal components (Fig. 3). (B) Boxplots of morphological variables of East Asian *A. donax* (white triangles,  $N = 26$ ) and eastern (white circles,  $N = 29$ ), northern (grey circles,  $N = 7$ ) and western (black circles,  $N = 23$ ) morphotypes of *A. formosana*, that is lengths (mm) of the lower glume (G1), the upper glume (G2), differences in the length between the lower and the upper glumes (G1–G2), lengths (mm) of the lemma (L) and the lemma awn (A), ratio of the awn length to the lemma length (A/L), lengths (mm) of the palea (pa) and the lemma hair (pL), number of flowers per spikelet (nbF), length of the stomatal guard cells (X,  $\mu\text{m}$ ), stoma density (dX, per  $10^4 \mu\text{m}^2$ ) and number of rib prickles per  $\mu\text{m}$  (dP, per mm of rib line). Continuous arrows indicate floral variables, dashed arrows indicate epidermis variables, grey values indicate an  $R^2$  coefficient of determination between East Asian *A. donax* and *A. formosana* taxa (and only in *A. formosana*) and crosses indicate values recorded from type specimens.



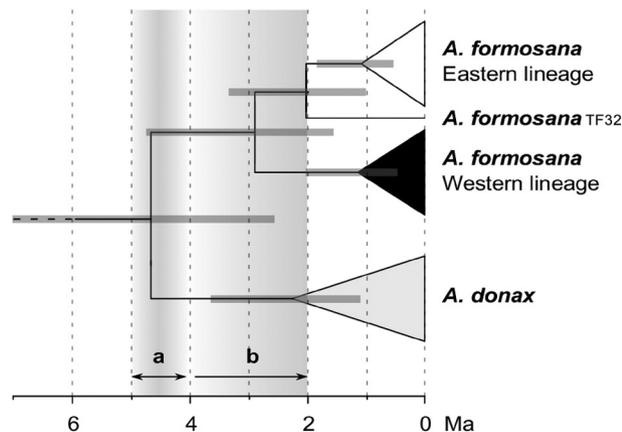
**Figure 3.** Geographical distribution of morphotypes (A, B), haplotypes (C, D) and genotypes (E, F) of *Arundo formosana* in Taiwan. (A) Morphometric assignment probabilities for the 61 samples (using a discriminant analysis) to the clusters defined with a (B) hierarchical clustering of principal components (designating  $k = 3$  as the best clustering of morphometric variation). Numbers indicate the location of type specimens for (1) *A. formosana*, (2) *A. formosana* var.

#### PLASTID DNA ANALYSES

The five plastid DNA sequences cover 4491 bp among 26 samples of East Asian *A. donax* and 33 of *A. formosana*. Without indels and outgroup considerations, this alignment includes 71 variable sites of which 46 are potentially parsimony-informative. *A. donax* and *A. formosana* are strongly supported in both MP (bootstrap, 72 and 92%, respectively; Fig. S2) and BI analyses [posterior probability (pp), 1 and 1; Fig. 4]. Concerning chronogram dating and according to JMODELTEST, *trnCF-rpoB* and *psaA-ORF170* sequences follow the HKY site model (unequal base frequencies and unequal rates of transitions and transversions) and *trnT-trnL*, *rbcL-psaI* and *trnS-psbD* the TPM1uf site model (unequal base frequencies and differences between one rate of transitions and two rates of transversions). The divergence between the two *Arundo* spp. is estimated to 4.66 Ma (95% confidence interval, 2.56–7.34 Ma) and the first divergence between the two main lineages of *A. formosana* to 2.97 Ma (95% confidence interval, 1.55–4.75 Ma; Fig. 4). These east and west lineages are well-supported by MP (bootstrap, 96 and 99%, respectively) and BI analyses (pp, 1 and 1).

The overall alignment of the five plastid DNA sequences for the 36 samples of *A. formosana* included 4433 bp, 60 variable sites, of which 14 were hypervariable (i.e. mini- and microsatellites), and 27 different haplotypes. Separated by 21 missing haplotypes and 22 mutation steps, the two main lineages clearly show an east–west geographical distinction. Only one sample/haplotype (TF32) from south–east Taiwan contradicts this two-directional pattern, with an early divergence from the eastern lineage (Fig. 4). Represented by 19 samples and 17 haplotypes, this eastern lineage is radially structured around one sample/haplotype, linked to 12 other haplotypes. Among them, one gives birth to a derived cluster including the four samples corresponding to the northern morphotype (Fig. 3C, D) plus another sample (TF70, central-eastern Taiwan). Including 15 samples and ten haplotypes, the western

*gracilis*, (3), *A. formosana* var. *robusta* and (4) *A. parviflora* (synonym *A. formosana*). (C) Geographical distribution and (D) phylogenetic tree of 27 plastid DNA haplotypes from 36 samples of *A. formosana* using a median-joining network algorithm (Network v.4.6; clustcolors arbitrarily attributed). (F) Analytic outputs assessing the best genetic clustering among AFLP data following the Delta  $K$  index (Evanno *et al.*, 2005) and (E) genotype composition of each sample ( $N = 28$ ) according to reconstructed ancestral lineages (STRUCTURE v.2.3). White circles tend to indicate *A. formosana* var. *formosana*, grey circles var. *robusta* and black circles subsp. *gracilis*. Topographic shading indicates elevations above 0, 1000 and 2000 m.



**Figure 4.** Chronogram of the genus *Arundo*. All nodal posterior probabilities = 1. (A) Emergence of the Taiwan Island (4–5 Ma). (B) Orogenesis of the Centre Mountain Range, reaching its current form 2 Ma ago. The divergence with *Phragmites australis* (Cav.) Steud. (outgroup, 20.9 Ma) is not shown. The eastern lineage represents *A. formosana* subsp. *formosana*, including var. *robusta*, and the western lineage represents *A. formosana* subsp. *gracilis*.

lineage is also structured around a central haplotype shared by six samples and linked to six other haplotypes, of which one gathers the three south-western samples/haplotypes. This western lineage also includes the DNA-preserved sample from Ryukyu Islands (K3; Figs 1 and 3C).

#### AFLPs

The two primer combinations generated 713 variable and repeatable markers among East Asian *A. donax* ( $N = 23$ ) and *A. formosana* ( $N = 36$ ), without deletions of DNA-degraded samples. The NJ tree generated from this dataset strongly supports the monophyly of each species (bootstrap, 100 and 100%). The minimization of BIC criterion in the DAPC analysis suggests these two species as the most suitable clustering of genetic diversity (Fig. S3).

In *A. formosana*, the RawGeno visualization of fragment number per sample led us to conserve 28 samples among the 36 samples analysed for plastid DNA sequences. After removal of non-replicable fragments, the two primer pairs generated 582 variable markers. The NJ tree generated in *A. formosana* also described eastern and western lineages (bootstrap, 92 and 96, respectively), representing the most suitable clustering following both the delta K statistic of structure analysis (Fig. 3E, F) and the BIC minimization of DAPC analysis (data not shown). Three samples from central-eastern Taiwan are located on the basis of this east–west bifurcation (TF23, 39 and 44; Fig. 5). However, these samples are designated as

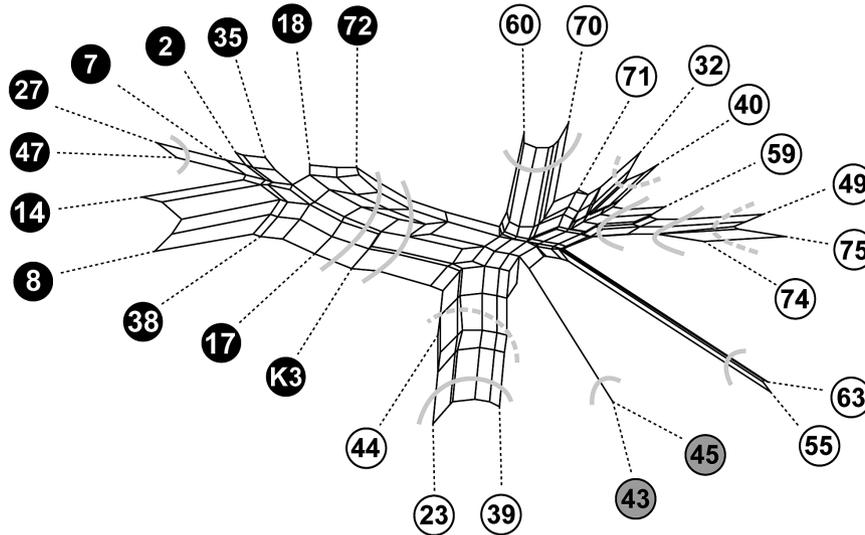
admixed genotypes between the two main lineages in STRUCTURE outputs (Fig. 3E). The two samples of the northern morphotype (TF43 and 45) successfully analysed in AFLPs were closely linked in the NJ tree (bootstrap, 100%) and NeighborNet (Fig. 5), despite their geographical distance. However, they do not form a major genetic cluster on the same level as eastern and western lineages, but rather an early divergence from the eastern lineage. The isolated sample from Ryukyu Islands (K3) is again assigned to the western lineage, despite the occurrence of a slight fraction of the eastern cluster in its reassigned genotype.

## DISCUSSION

### MORPHOLOGICAL DEFINITION OF VARIETIES

When describing *A. formosana*, Hackel (1899) also described *A. formosana* var. *gracilis*, a smaller and more branched phenotype distinguished from the type variety *A. formosana* var. *formosana* by its leaf blades 10 (vs. 20) cm long and 0.5 (vs. 1.0) cm wide, panicles 10 (vs. 35) cm long, spikelet 4 (vs. 5) mm long and lemma awn 1.5 (vs. 2.5) mm long. Our data suggest that leaf dimensions are not in fact the most suitable characters for species identification due to their high morphological variability and the likely biased selection of material preserved on herbarium specimens. Our data do, however, show support for Hackel's variety with distinctly different measurements of flowering parts. Much later, Conert (1961) added to Hackel's classification by describing another variety, *A. formosana* var. *robusta*, distinguished on flowering parts mainly by a long lemma awn representing more than the half of the lemma (awn, 4.0–4.5 mm; lemma, 6.5–7.0 mm). Our study is based on broader sampling than anything available to Hackel or to Conert and employs statistical analyses of morphometric data. Our results also divide *A. formosana* into three morphotypes, each one including a type specimen (indicated with crosses in Figs 2 and 6). Despite this morphometric congruence, we did not succeed in documenting further qualitative and non-overlapping characters to distinguish these morphotypes. A weak aspect of our study is the record of pubescence cover which was limited by its variable preservation on herbarium specimens. The critical importance of indumentum for taxonomy of *Arundo* has been demonstrated by Hardion et al. (2012a), and further investigation of the indumentum is needed with fresh material. We use our morphometric dataset to update Conert's key distinguishing between the three morphotypes, with the addition of the East Asian *A. donax* (see Taxonomic Treatment section).

In the taxonomic literature, description of subspecies and varieties is usually supported by overlapping



**Figure 5.** NeighborNet based on 28 AFLP fingerprints of *A. formosana*. Sample colors indicate main genotype assignment based on STRUCTURE analysis (Figs 3E, F). Dashed and continuous grey lines indicate the position of nodes from the neighbor-joining tree with bootstrap values superior to 50 and 75%, respectively. White circles indicate *A. formosana* var. *formosana*, grey circles var. *robusta* and black circles subsp. *gracilis*.

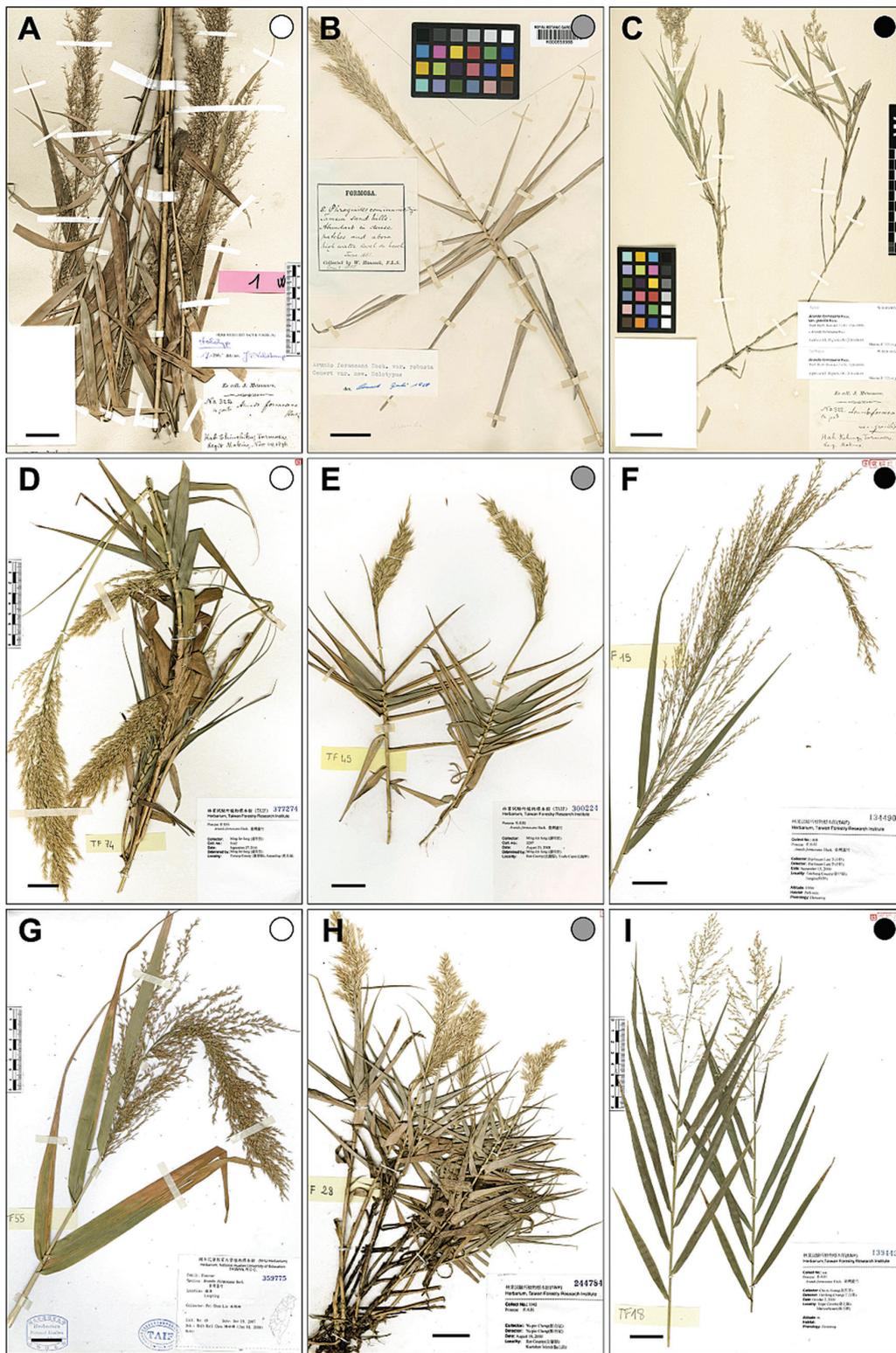
sets of morphological characters. Our overlapping measurements for the three varieties of *A. formosana* match this standard practice of infraspecific taxonomy. The morphometric boxplots and our identification key offer a reliable distinction between these infraspecific units. The geographical distribution of these morphotypes is quite clearly structured, with eastern, western and northern morphotypes (Fig. 3A). Type specimens also follow this geographical pattern, with one exception: the western origin of the type collection of var. *formosana*. This mismatch between morphology and geography could be due to incorrect determination of the type specimen, easily possible considering the overlap in floral measurements between the eastern and western morphotypes. However, this uncertainty cannot be resolved by standard molecular analysis because the type specimens of *A. formosana* varieties were collected between 1881 and 1896 (Fig. 6).

#### PHYLOGEOGRAPHY OF *A. FORMOSANA* IN TAIWAN

The island of Taiwan is a diversity hotspot for *Arundo* and for much Asian biodiversity, with 25% endemism in a flora of over 4000 plant species (Hsieh, 2002; Chiang & Schaal, 2006). The island started to separate from the mainland in only the recent geological past, 4–5 Ma (early Pliocene; Shaw, 1996; Lin & Watts, 2002), gaining its current position of 150 km from south-eastern China (Fig. 1). Ice cover during Pleistocene glaciations is thought to have been incomplete allowing the conservation of temperate to subtropical taxa in several refugia (Wu *et al.*, 2006a). Despite its elongated

shape from north to south (394 km long vs. 144 km wide), the diversity distribution of Taiwan is structured primarily east to west due to the latitudinal orogeny of a central mountain range reaching its present shape at *c.* 2 Ma (Huang, Yuan & Tsao, 2006). This central mountain range has acted as a barrier for most of the 25 plant and animal taxa studied in Taiwan-wide phylogeographic evaluations (e.g. Toda *et al.*, 1998; Wang, Hsu & Chiang, 2000; Huang, Hwang & Lin, 2002; Cheng, Hwang & Lin, 2005; Oshida *et al.*, 2006; Wu *et al.*, 2006a, 2007; Jang-Liaw, Lee & Chou, 2008; Kuo *et al.*, 2014; Yu, Lin & Weng, 2014).

The evolutionary history of *A. formosana* is congruent with the recent orogenesis on Taiwan. Molecular dating indicates that the split between the Taiwanese *A. formosana* and East Asian *A. donax* corresponds to the first emergence of this island estimated at 4–5 Ma (Huang *et al.*, 2006). After the formation of Taiwan, these species appear as two independent monophyletic lineages supported by both plastid and nuclear data (Figs S2 and S3). The current occurrence of *A. donax* in Taiwan could be due to later natural or human dispersal events, this weed species now being both an ornamental and an invasive plant. Between 4 and 2 Ma, formation of the central mountain range (CMR) created a barrier between populations of *A. formosana* and led to the divergence of two infraspecific lineages around 3 Ma. A third much differentiated haplotype, only represented by sample TF32, may indicate persistence of rare haplotypes from early divergences in the southern part of the island, even though this deep differentiation was not supported by



**Figure 6.** Herbarium specimens of *Arundo formosana* var. *formosana* (white circles) from Shinchiku (A; holotype; W1916-0034625), Aopanling (D; TF74) and Lingting (G; TF55); *A. formosana* var. *robusta* (grey circles) from Tamsui (B; holotype; K000859988), Toufu Cape (E; TF45) and Kuan-in Shan (H; TF28); and *A. formosana* subsp. *gracilis* (black circles) from Kelung (C; W1916-0034626), Tunplin (F; TF15) and Manyuehyuan (I; TF18). Scale bar, 3 cm.

AFLP data. Genomic divergence shown by the AFLP dataset clearly supports the east–west division of the two main lineages. The admixed AFLP genotypes also suggest some secondary contact and gene flows in the valleys of the northern part of the CMR for this wind-pollinated and wind-dispersed species. Eastern samples from the Ryukyu Islands seem more related to the western lineage (Figs 3B, C).

#### CONSENSUS CLASSIFICATION

Intraspecific taxonomy has historically been used to describe allopatric groups of populations with weak morphological distinctions (Davis & Heywood, 1963; Mayr, 1963). Taxonomists have also stressed the ecological differentiation of infraspecific ranks, sometimes treating them as equivalent to ecotypes (Venu, 2002; Slovák *et al.*, 2012). Evidence of genetic divergence within understudied taxa has stimulated testing of infraspecific structure based on morphological features (e.g. in *Phragmites australis*; Saltonstall, Peterson & Soreng, 2004). This study uses a pre-existing, but neglected infraspecific classification of *A. formosana* to test the phylogenetic divergence between its varieties.

Our statistical exploration of morphometric variation confirmed the distinctiveness of three geographically separate entities, each one containing the type collection of a single variety. The three morphotypes were taken as the first taxonomic hypothesis to test with plastid DNA and AFLP data which largely support the western var. *gracilis* and the eastern var. *formosana* as the main lineages in *A. formosana*. The *robusta*-morphotype accessions form a single phylogenetic group based on plastid DNA and a distinct group based on AFLPs; this variety represents a secondary smaller divergence in the evolutionary history

of *A. formosana*. Consequently, a strict consensus of the three markers based on main evolutionary lineages ought to recognize only two infraspecific entities in *A. formosana*. However, the obvious morphological differentiation of the *robusta*-morphotype (Figs 2 and 7) argues in favour of the recognition of this variety. In addition, molecular data still highlight a secondary genetic coherence of var. *robusta*, despite its paraphyletic definition based on the plastid DNA and a weak sampling on AFLPs (only two samples analysed; Fig. 4). Unlike the two major lineages, the differentiation of var. *robusta* cannot be explained solely by geographical isolation and phenotypic drift. The longer reproductive structures of var. *robusta* could be linked to polyploid ‘gigantism’ observed for other *Arundo* spp. (Hardion *et al.*, 2012a, 2014a, b, 2015). Finally, our genome-wide AFLP data do not support the differentiation of var. *robusta* as the result of a hybridization between the two main genetic lineages previously described. If var. *robusta* can be included in the eastern phylogenetic lineage, its northern distribution close to locations of the western lineage does not reveal admixed genotypes. For this wind-pollinated and wind-dispersed species, the lack of genetic admixture in the north could be explained by a biological (or at least ecological) barrier isolating var. *robusta* genotypes, but our sampling remains still weak to test this hypothesis.

#### SPECIES, SUBSPECIES OR VARIETIES?

In this study we have generated the first genetic and morphometric datasets based on a broad sampling of *A. formosana*, allowing an update of its taxonomy. To update the taxonomy, a choice of the appropriate taxonomic rank must be made for each subdivision. To define distinctiveness at the specific and infraspecific



**Figure 7.** (A) *Arundo formosana* subsp. *gracilis* located in a humid cliff from Iriomote, with long decumbent culms (©T. Takamiso); (B, C) *A. formosana* var. *robusta* on the coast of Keelung Yehliu Geopark, with short erect culms (©Xu Yechun).

ranks, we repeated every analysis on a larger dataset including East Asian *A. donax*. Every clustering analysis supports clear reciprocal monophyly of *A. formosana* and *A. donax*, confirming the application of species rank for these two taxa. The definition of clusters in *A. formosana* was only possible in a species-specific analysis, that is after the removal of *A. donax* samples from the dataset. To define species and infraspecific ranks, some taxonomists apply the biological definition of a species, assuming the possibility of crossing between infraspecific ranks in a sympatric zone (e.g. Yoshino *et al.*, 2011). Our AFLP data support this with the evidence of a putative hybrid zone in the northern part of the CMR between the two main genetic lineages of *A. formosana*. The choice between the subspecies and the variety rank was widely debated in the mid-20th century (Clausen, 1941; Fosberg, 1942; Weatherby, 1942; Kapadia, 1963; Raven, 1974). We have chosen to follow articles 4.1 and 4.2 of the ICBN (McNeil *et al.*, 2012), stipulating that variety is the prevailing rank below the species and reserving the use of subspecies for the clustering of varieties. Because our results clearly describe two main genetic lineages as geographically and morphologically distinct, of which one includes a secondary genetic cluster corresponding to an ecotype, we choose to accept two subspecies, one of which includes a variety. We recognize *A. formosana* var. *robusta* and *A. formosana* var. *formosana* in the type subspecies *A. formosana* subsp. *formosana*. We also recognize a second subspecies *A. formosana* subsp. *gracilis* (Hack.) Hardion, Verlaque & B. Vila.

#### TAXONOMIC TREATMENT

***Arundo formosana*** Hack. in Bull. Herb. Boiss. 7: 724. 1899 – Type: Taiwan. Prope Shinchiku, 24 November 1896, Makino 322 (holotype: W!; isotype: US!).

= *Arundo parviflora* Ohwi in Repertorium Specierum Novarum Regni Vegetabilis 36: 40. 1934 – Type: Japan (Taiwan under Japanese rule). prov. Takaoshu, inter Matsuyama et Aderu in Chippingoe, 7 May 1933, Ohwi 1597 (holotype: KYO! photograph; isotype: US!).

***Arundo formosana* subsp. *formosana***

***Arundo formosana* subsp. *formosana* var. *formosana***

***Arundo formosana* subsp. *formosana* var. *robusta*** Conert in Die Systematik und Anatomie der Arundineae: 35. 1961 – Type: Taiwan. Tamsui, sand hills, June 1881, Hancock 8 (holotype: K!).

***Arundo formosana* subsp. *gracilis*** (Hack.) Hardion, Verlaque & B. Vila, comb. nov.

≡ *Arundo formosana* var. *gracilis* Hack. in Bull. Herb. Boiss. 7: 724. 1899 – Type: Taiwan. Prope Kelung, Makino 322 (holotype: W!; isotype: US!).

#### DETERMINATION KEY

1. Glumes 7–12 mm long, seeming equal on spikelet, the lower slightly longer than the upper, palea 4–6 mm long, lemma hairs 3–6 mm long, lemma 7–12 mm with awn 1–3 mm long.....***A. donax***
1. Glumes 3–9 mm long, seeming unequal on spikelet, the lower slightly shorter than the upper, palea 2–4 mm long, lemma hairs 1–4 mm long, lemma 4–12 mm with awn 0–6 mm long...***A. formosana*: 2**
2. Lemma 6–12 mm with awn 1–6 mm long, glumes 4–9 mm long.....subsp. ***formosana*: 3**
2. Lemma 4–6 mm with awn 0–3 mm long, glumes 3–6 mm long.....subsp. ***gracilis***
3. Spikelet one- to three-flowered, glumes 4–7 mm long, lemma 6–10 mm long with awn 1–5 mm long .....var. ***formosana***
3. Spikelet three- to six-flowered, glumes 7–9 mm long, lemma 8–12 mm long with awn 2–6 mm long .....var. ***robusta***

#### CONCLUSION

Intraspecific ranks are undervalued in phylogeographic studies and an integrative approach could reconcile the practical aspects of infraspecific taxonomy, that is the definition of formal units based on diagnostic characters, with the genetic data and the evolutionary framework provided by molecular markers. Varieties and subspecies could be treated as infraspecific groups which are statistically distinguishable on morphological characters, despite having overlapping ranges, and with a geographical, ecological and/or reproductive isolation supported by genetic markers (including paraphyletic groups). Phylogeography could be one of the most suitable approaches for the investigation of evolutionary histories of subspecies and varieties and phylogeographic studies should take such taxonomic issues into consideration.

#### ACKNOWLEDGEMENTS

The authors thank Françoise Fleury-Blanc (MARS, Aix-Marseille University), Shih-Wen Chung (TAIF, Taiwan Forestry Research Institute), Ernst Vitek (W, NHM Vienna), Martin Xanthos (K, RBG Kew) and Andrew Clark and Paul Peterson (US, Smithsonian Institute) for the management of herbarium loans; Julie Guery and Quentin Delforge for morphometrical measures; Marianick Juin and Nathalie Duong (IMBE) for molecular work; Anne Roig and Bruno Fady (INRA) for the generation of AFLP electroforetograms; and Arne Saatkamp (IMBE) for the German-to-English translation of Conert's determination keys.

## REFERENCES

- Arrigo N, Tuszynski JW, Ehrich D, Gerdes T, Alvarez N. 2009.** Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *BMC Bioinformatics* **10**: 33.
- Bandelt HJ, Forster P, Röhl A. 1999.** Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**: 37–48.
- Braby MF, Eastwood R, Murray N. 2012.** The subspecies concept in butterflies: has its application in taxonomy and conservation biology outlived its usefulness? *Biological Journal of Linnean Society* **106**: 699–716.
- Cheng YP, Hwang SY, Lin TP. 2005.** Potential refugia in Taiwan revealed by the phylogeographical study of *Castanopsis carlesii* Hayata (Fagaceae). *Molecular Ecology* **14**: 2075–2085.
- Chiang TY, Schaal BA. 2006.** Phylogeography of plants in Taiwan and the Ryukyu Archipelago. *Taxon* **55**: 31–41.
- Christin PA, Spriggs E, Osborne CP, Strömberg CA, Salamin N, Edwards EJ. 2014.** Molecular dating, evolutionary rates, and the age of the grasses. *Systematic Biology* **63**: 153–165.
- Clausen RT. 1941.** On the use of the terms ‘subspecies’ and ‘variety’. *Rhodora* **43**: 157–167.
- Conert HJ. 1961.** *Die Systematik und Anatomie der Arundinae*. Weinheim: Verlag von J. Cramer.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012.** jModel-Test 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772.
- Davis PH, Heywood VH. 1963.** *Principles of angiosperm taxonomy*. Edinburgh: Oliver & Boyd.
- De Queiroz K. 2007.** Species concepts and species delimitation. *Systematic Biology* **56**: 879–886.
- Doyle JJ, Doyle JL. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19**: 11–15.
- Dray S, Dufour AB. 2007.** The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* **22**: 1–20.
- Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006.** Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**: e88.
- Ehrich D. 2007.** *Structure-sum v. 2007: a series of R functions for summarizing the outputs of the program Structure ver. 2.2*. Unpublished, available from the author upon request.
- Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.
- Fosberg FR. 1942.** Subspecies and variety. *Rhodora* **44**: 153–157.
- Govaerts R. 1999.** *World checklist of seed plants, vol. 3, pt. 1*. Antwerp: Continental Publishing.
- Hackel E. 1899.** Enumeratio graminum Japoniae. Verzeichnis der Gräser Japans. *Bulletin de l’Herbier Boissier* **7**: 701–726.
- Hamilton CW, Reichard SH. 1992.** Current practice in the use of subspecies, variety, and forma in the classification of wild plants. *Taxon* **41**: 485–498.
- Hardion L, Baumel A, Verlaque R, Vila B. 2014a.** Distinct evolutionary histories of lowland biota on Italian and Balkan peninsulas revealed by the phylogeography of *Arundo plinii* (Poaceae). *Journal of Biogeography* **41**: 2150–2161.
- Hardion L, Verlaque R, Baumel A, Juin M, Vila B. 2012a.** Revised systematics of Mediterranean *Arundo* (Poaceae) based on AFLP fingerprints and morphology. *Taxon* **61**: 1217–1226.
- Hardion L, Verlaque R, Callmander MW, Vila B. 2012b.** *Arundo micrantha* Lam. (Poaceae), the correct name for *Arundo mauritanica* Desf. and *Arundo mediterranea* Danin. *Candollea* **67**: 131–135.
- Hardion L, Verlaque R, Rosato M, Rosselló JA, Vila B. 2015.** Impact of polyploidy on fertility variation of Mediterranean *Arundo* L. (Poaceae). *Comptes rendus biologies* **338**: 298–306.
- Hardion L, Verlaque R, Saltonstall K, Leriche A, Vila B. 2014b.** Origin of the invasive *Arundo donax* (Poaceae): a trans-Asian expedition in herbaria. *Annals of Botany* **114**: 455–462.
- Hilu KW, Randall JL. 1984.** Convenient method for studying grass leaf epidermis. *Taxon* **33**: 413–415.
- Hsieh CF. 2002.** Composition, endemism and phytogeographical affinities of the Taiwan flora. *Taiwania* **47**: 298–310.
- Huang CY, Yuan PB, Tsao SJ. 2006.** Temporal and spatial records of active arc-continent collision in Taiwan. *Geological Society of America Bulletin* **118**: 274–288.
- Huang SS, Hwang SY, Lin TP. 2002.** Spatial pattern of chloroplast DNA variation of *Cyclobalanopsis glauca* in Taiwan and East Asia. *Molecular Ecology* **11**: 2349–2358.
- Huang TC, Boufford DE, Hsieh CF, Kuoh CS, Ohashi H, Su HJ. 2000.** *Flora of Taiwan*. Taipei: National Taiwan University.
- Huson DH, Bryant D. 2006.** Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**: 254–267.
- Husson F, Josse J, Pages J. 2010.** Principal component methods-hierarchical clustering-partitional clustering: why would we need to choose for visualizing data. *Technical Report – Agrocampus, Applied Mathematics Department*.
- Jang-Liaw NH, Lee TH, Chou WH. 2008.** Phylogeography of *Sylvirana latouchii* (Anura, Ranidae) in Taiwan. *Zoological Science* **25**: 68–79.
- Jombart T, Devillard S, Balloux F. 2010.** Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**: 94.
- Kapadia ZJ. 1963.** Varietas and subspecies, a suggestion towards greater uniformity. *Taxon* **12**: 257–259.
- Kuo HC, Chen SF, Fang YP, Flanders J, Rossiter SJ. 2014.** Comparative range wide phylogeography of four endemic Taiwanese bat species. *Molecular Ecology* **23**: 3566–3586.
- Lin AT, Watts AB. 2002.** Origin of the West Taiwan basin by orogenic loading and flexure of a rifted continental margin. *Journal of Geophysical Research-Solid Earth* **107**: B9, 2185.

- Mayr E. 1963. *Animal species and evolution*. Cambridge: Belknap Press of Harvard University Press.
- McNeil J, Barrie FR, Buck WR, Demoulin V, Greuter W, Hawksworth DL, Herendeen PS, Knapp S, Marhold K, Prado J, Prud'homme van Reine WF, Smith GF, Wiersema JH, Turland NJ. 2012. *International code of nomenclature for algae, fungi, and plants (Melbourne Code)*. *Regnum Vegetabile* 154. Königstein: Koelz Scientific Books.
- Moritz C. 1994. Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology & Evolution* 9: 373–375.
- O'Brien SJ, Mayr E. 1991. Bureaucratic mischief: recognizing endangered species and subspecies. *Science* 251: 1187–1188.
- Oshida T, Lee JK, Lin LK, Chen YJ. 2006. Phylogeography of Pallas's squirrel in Taiwan: geographical isolation in an arboreal small mammal. *Journal of Mammalogy* 87: 247–254.
- Phillimore AB, Owens IP. 2006. Are subspecies useful in evolutionary and conservation biology? *Proceedings of the Royal Society B* 273: 1049–1053.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- R Development Core Team. 2015. *R: a language and environment for statistical computing*. Vienna: R Foundation for statistical computing.
- Rambaut A, Drummond AJ. 2007. *Tracer v.1.4. Computer program and documentation distributed by the authors*. Available at: <http://beast.bio.ed.ac.uk/Tracer>.
- Raven PH. 1974. Nomenclature proposals to the Leningrad congress. *Taxon* 23: 828–833.
- Saltonstall K. 2001. A set of primers for amplification of non-coding regions of chloroplast DNA in the grasses. *Molecular Ecology Notes* 1: 76–78.
- Saltonstall K, Lambertini C. 2012. The value of repetitive sequences in chloroplast DNA for phylogeographic inference: a comment on Vachon & Freeland 2011. *Molecular Ecology Resources* 12: 581–585.
- Saltonstall K, Peterson PM, Soreng RJ. 2004. Recognition of *Phragmites australis* subsp. *americanus* (Poaceae: Arundinoideae) in North America: evidence from morphological and genetic analyses. *SIDA, Contributions to Botany* 21: 683–692.
- Shaw CL. 1996. Tectonic evolution of the Cretaceous-Cenozoic basins of Taiwan. *Annual Report of National Taiwan Museum* 39: 429–446.
- Sites JW, Marshall JC. 2003. Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology and Evolution* 18: 462–470.
- Sites JW Jr, Marshall JC. 2004. Operational criteria for delimiting species. *Annual Review of Ecology, Evolution and Systematics* 35: 199–227.
- Slovák M, Kucera J, Marhold K, Zozomová-Lihová J. 2012. The morphological and genetic variation in the polymorphic species *Picris hieracioides* (Compositae, Lactuceae) in Europe strongly contrasts with traditional taxonomical concepts. *Systematic Botany* 37: 258–278.
- Swofford DL. 2003. *PAUP\*: phylogenetic analysis using parsimony (\* and other methods), version 4.0*. Sunderland: Sinauer Associates.
- Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731–2739.
- Toda M, Nishida M, Matsui M, Lue KY, Ota H. 1998. Genetic variation in the Indian rice frog, *Rana limnocharis* (Amphibia: Anura), in Taiwan, as revealed by allozyme data. *Herpetologica* 54: 73–82.
- Venables WN, Ripley BD. 2002. *Modern applied statistics with S*. Berlin: Springer.
- Venu P. 2002. Some conceptual and practical issues in taxonomic research. *Current Science* 82: 924–933.
- Vorontsova MS, Ratovonirina G, Randriamboavonjy T. 2013. Revision of *Andropogon* and *Diectomis* (Poaceae: Sacchareae) in Madagascar and the new *Andropogon itremoensis* from the Itremo Massif. *Kew Bulletin* 68: 193–207.
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.
- Wang JP, Hsu KC, Chiang TY. 2000. Mitochondrial DNA phylogeography of *Acrossocheilus paradoxus* (Cyprinidae) in Taiwan. *Molecular Ecology* 9: 1483–1494.
- Weatherby CA. 1942. Subspecies. *Rhodora* 44: 157–167.
- Wilson JJ, Sing KW, Sofian-Azirun M. 2013. Building a DNA barcode reference library for the true butterflies (Lepidoptera) of Peninsula Malaysia: what about the subspecies? *PLoS One* 8: e79969.
- Wu JH, Hsu CH, Fang LS, Chen IS. 2007. The molecular phylogeography of *Candidia barbata* species complex (Teleostei: Cyprinidae) from Taiwan. *Raffles Bulletin of Zoology Supplements* 14: 61–67.
- Wu SH, Hwang CY, Lin TP, Chung JD, Cheng YP, Hwang SY. 2006a. Contrasting phylogeographical patterns of two closely related species, *Machilus thunbergii* and *Machilus kusanoi* (Lauraceae), in Taiwan. *Journal of Biogeography* 33: 936–947.
- Wu ZY, Raven PH, Hong DY. 2006b. *Flora of China: Poaceae, Vol. 22*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press.
- Yoshino N, Wang GX, Uchino A, Tominaga T. 2011. Hybridizations and genetic relationships among *Lindernia* species (Scrophulariaceae): *L. procumbens* and two subspecies of *L. dubia*. *Aquatic Botany* 94: 165–171.
- Yu TL, Lin HD, Weng CF. 2014. A new phylogeographic pattern of endemic *Bufo bankorensis* in Taiwan Island is attributed to the genetic variation of populations. *PLoS One* 9: e98029.
- Zink RM. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proceedings of the Royal Society B* 271: 561–564.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website:

**Table S1.** Sampling information, morphotype assignment (M), AFLP phylogenetic lineage (G), haplotype combinations (H) and GenBank accessions for *psaA*-ORF170 (AO), *trnCF-rpoB* (CB), *rbcL-psaI* (LI), *trnS(GCU)-psbD* (SD) and *trnT-trnL* (TL).

**Figure S1.** Hierarchical clustering of principal components based on morphometric variables, validating the species level between *Arundo formosana* and East Asian *A. donax* ( $k = 2$ ) as the best clustering of morphometric variation.

**Figure S2.** Molecular phylogenetic tree for *Arundo formosana* and East Asian *A. donax* using maximum parsimony and the nucleotide variation of five plastid DNA spacers (value indicated bootstrap value among 1000 replicates).

**Figure S3.** Neighbor-joining tree based on AFLP data among *Arundo formosana* and East Asian *A. donax*. Black, grey and absence of triangle indicate bootstrap value  $>75$ ,  $>50$  and  $<50\%$ , respectively (based on 1000 bootstrap replicates).