



SPECIAL  
PAPER



# An evaluation of the robustness of global amphibian range maps

Gentile Francesco Ficetola<sup>1\*</sup>, Carlo Rondinini<sup>2</sup>, Anna Bonardi<sup>1</sup>, Vineet Katariya<sup>3</sup>, Emilio Padoa-Schioppa<sup>1</sup> and Ariadne Angulo<sup>3</sup>

<sup>1</sup>Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, 20126, Milano, Italy, <sup>2</sup>Global Mammal Assessment program, Department of Biology and Biotechnologies, Sapienza University of Rome, Viale dell'Università 32, I-00185, Roma, Italy, <sup>3</sup>IUCN Global Species Programme, Rue Mauverney 28, 1196, Gland, Switzerland

## ABSTRACT

**Aim** Maps of species ranges are among the most frequently used distribution data in biodiversity studies. As with any biological data, range maps have some level of measurement error, but this error is rarely quantified. We assessed the error associated with amphibian range maps by comparing them with point locality data.

**Location** Global.

**Methods** The maps published by the Global Amphibian Assessment were assessed against two data sets of species point localities: the Global Biodiversity Information Facility (GBIF), and a refined data set including recently published, high-quality presence data from both GBIF and other sources. Range fit was measured as the proportion of presence records falling within the range polygon(s) for each species.

**Results** Using the high-quality point data provided better fit measures than using the raw GBIF data. Range fit was highly variable among continents, being highest for North American and European species (a fit of 84–94%), and lowest for Asian and South American species (a fit of 57–64%). At the global scale, 95% of amphibian point records were inside the ranges published in maps, or within 31 km of the range edge. However, differences among continents were striking, and more points were found far from range edges for South American and Asian species.

**Main conclusions** The Global Amphibian Assessment range maps represent the known distribution of most amphibians well; this study provides measures of accuracy that can be useful for future research using amphibian maps as baseline data. Nevertheless, there is a need for greater investment in the continuous updating and improvement of maps, particularly in the megadiverse areas of tropical Asia and South America.

## Keywords

Alpha-hulls, amphibians, data quality, conservation biogeography, GBIF, measurement error, point data, range size, species distribution range, Wallacean shortfall.

\*Correspondence: Gentile Francesco Ficetola, Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, 20126 Milano, Italy. E-mail: francesco.ficetola@unimib.it

## INTRODUCTION

In recent years, a growing number of studies have investigated patterns of biodiversity at broad spatial scales. These studies are helping us to understand the factors determining species distribution, richness and abundance, thus providing the information needed to set up conservation strategies (Grenyer *et al.*, 2006; Finer *et al.*, 2008; Lawler *et al.*, 2010; Rondinini

*et al.*, 2011; Sandel *et al.*, 2011; Hof *et al.*, 2012; Ficetola *et al.*, 2013). Direct field sampling over very large spatial scales is rarely feasible, as it would require significant resources and time. Broad-scale biodiversity analyses must therefore rely on a variety of species distribution data, ranging from point locality data obtained from databases, faunistic atlases, geographical range maps and species distribution models. Unfortunately, our knowledge of biodiversity distribution is

far from complete, and we have a limited knowledge of actual species distribution even for the best-studied taxa (Lomolino, 2004; Mokany & Ferrier, 2011; Ficetola *et al.*, 2013).

The quality of broad-scale biodiversity analyses and their usefulness for conservation purposes is strongly related to the data on which analyses are based, but any biological data have some level of measurement error. Among species distribution data, errors are routinely quantified for some data types (species distribution models) but not for others (point localities, geographical range maps) (Rondinini *et al.*, 2006; Rocchini *et al.*, 2011). Geographical range maps encompass the broad areas where a species is thought to be found, and assume the species' presence inside the range and absence outside. Even with this assumption, tests are needed to estimate the reliability of the range edge (Gaston, 2003; Rocchini *et al.*, 2011). In the real world, the range edge is not always a hard boundary, except when it is represented by a physical barrier, e.g. the coastline for terrestrial species. With soft boundaries, populations at range margins can be sinks that experience extinction–recolonization dynamics and can be at lower densities than core populations (Gaston, 2003; Kanda *et al.*, 2009). This is the case, for example, with valleys for species living on mountain tops, or savanna for forest species, or with boundaries defined by more complex biotic interactions (e.g. competition or predation). The latter case may include the range edge of many species that live in major forest basins (Amazon and Congo) but do not occupy the basin entirely (Gaston, 2003; Naka *et al.*, 2012). As a consequence, true range boundaries often move over time and are difficult to assess (Gaston, 2003; Thomas, 2010). Range margins are therefore expected to be more accurately detected in regions with high sampling intensity, rather than regions where species knowledge is poor (one of the facets of the Wallacean shortfall, i.e. the uneven knowledge on species distribution across space) (Lomolino, 2004; Rocchini *et al.*, 2011).

The maps of amphibian species ranges, originally published by the Global Amphibian Assessment (GAA) in 2004 and housed in the IUCN Red List of Threatened Species (available at: <http://www.iucnredlist.org/>), are now a milestone for studies of amphibian ecology and conservation. They are the only maps of amphibian distribution available at the global scale, are routinely considered the most authoritative resource for large-scale amphibian distributions, and are used in an increasing number of studies on amphibian biogeography, ecology, evolution and conservation, as well as in bioclimatic species distribution models (e.g. Cooper *et al.*, 2008; Finer *et al.*, 2008; van Wilgen *et al.*, 2009; Lawler *et al.*, 2010; Becker & Zamudio, 2011; Sandel *et al.*, 2011; Hof *et al.*, 2012), including analyses performed on fine spatial scales (e.g. Finer *et al.*, 2008; van Wilgen *et al.*, 2009; Becker & Zamudio, 2011). The maps may be affected by multiple sources of error, such as incomplete information on some species or in some areas, limited spatial resolution, and errors when digitizing the distribution ranges, which may influence the output of analyses based on these maps (Hurlbert & Jetz, 2007; Foody, 2011; Rocchini *et al.*, 2011; Cantú-Salazar & Gaston, 2013).

Determining the level of accuracy across the spatial database can improve the use of IUCN amphibian data in ecology, conservation and evolutionary biology, and allows a better understanding of the strengths and limitations of analyses based on these maps (Hurlbert & Jetz, 2007; Rocchini *et al.*, 2011). It can also highlight those areas that may need greater investment in field surveys and data collation.

The aim of this study was to quantify the error associated with the edges of the range maps (hereafter referred to as 'amphibian ranges'). The quality of the data mostly depends on expert knowledge, distribution of surveyor effort and access to the field data, compared to the number of species to be mapped. Recent analyses showed strong geographical bias of terrestrial ecological observations, with a concentration of studies in rich continents with limited biodiversity (Europe and North America), while fewer studies have been performed in areas with high biodiversity but limited or more recent economic growth (particularly Africa and South America) (Martin *et al.*, 2012). Analogous geographical patterns are observed in the global distribution of amphibian taxonomists (Rodrigues *et al.*, 2010). We therefore predicted that the errors associated with the amphibian range maps parallel the global pattern of ecological observations, with smaller errors in the continents that have received more studies, compared to their biodiversity.

To achieve our goal, we compared the amphibian range maps with presence records of amphibian species obtained from the Global Biodiversity Information Facility (GBIF) and some of the recent primary literature. We assessed the fit between amphibian ranges and point locality data through two data sets: a very comprehensive but unverified point locality data set, comprising distribution records for the majority of species; and a smaller data set, considering only recent distribution records that are likely to be accurate. To this end, we defined 'range fit' as the proportion of presence records that can be found within the amphibian range polygon(s) for each species. As uncertainty may also be present in the validation data (i.e. point localities), we also tested how the quality of validation data may affect the validation process. Furthermore, we provide explicit measures of spatial accuracy of ranges by identifying confidence bands that may allow us to reduce the risk of omitting unmapped populations from analyses. The results of our study provide insights into the nature of maps and GBIF data, but also identify priorities for future research/work.

## MATERIALS AND METHODS

### Range maps

The first comprehensive conservation assessment of all amphibian species was completed in 2004 through the GAA (Stuart *et al.*, 2004; Chanson *et al.*, 2008a) following the taxonomy of the *Amphibian Species of the World* (Frost, 2004), and produced maps for all of those species, with a few (approximately 100) exceptions, primarily due to uncertain

geographical provenance and/or taxonomic uncertainty (Stuart *et al.*, 2008). Initial ranges were represented as polygons in spatial data files; for several species, polygons were drawn to encompass known point localities (Stuart *et al.*, 2008; N. Cox, IUCN Global Species Programme, pers. comm., January 2012). The regional coordinators in charge of initially developing local assessments drew these polygons either in the Species Information System Data Entry Module tool, a relatively simple program developed for assessment purposes, or using other spatial data software. The integration of expert assessment and workshops within a geographical data framework provides an excellent baseline for conservation evaluations (Sanderson *et al.*, 2002). Hard copies of these initial ranges were thus reviewed in GAA workshops by experts and modified by hand, with changes later incorporated into vector polygons (shapefiles). After the workshops, polygons were clipped to coastlines, and sometimes polygons were also clipped to rivers or elevational contours, although there was variation in the consistency with which this was done (Stuart *et al.*, 2008; N. Cox, pers. comm., January 2012). Of the 33 geographical units that were identified by the GAA coordinating team (e.g. Africa, Borneo, Caribbean, North America, etc., see Stuart *et al.*, 2008), the only region that was mapped using political boundaries was North America, where maps were developed by addition of the county limits where species were present (maps originated from the Amphibian Research and Monitoring Initiative, ARMI; Muths *et al.*, 2006; S. Stuart, IUCN Species Survival Commission, pers. comm., 2012). This could have occurred because US herpetological atlases usually use county limits as their spatial sampling units, in contrast with other areas of the world (e.g. Europe uses UTM squares). In general, greater attention was paid when mapping species with restricted ranges because they were of greater conservation concern. If possible, ranges were not extrapolated beyond known limits, even when wider ranges were considered likely. However, this was not the case for some areas; for instance, records existing on either side of the Congo Basin were taken to indicate that a species was present within the largely unexplored centre of the basin (N. Cox, pers. comm., 2012). Polygons were fragmented or adjusted for large areas of obviously unsuitable habitat (S. Stuart, pers. comm., 2012), but there was variation in the consistency with which this was applied owing to ambiguities in the IUCN guidelines at the time (N. Cox, pers. comm., 2012). Since 2004, smaller updates (primarily the addition of new species and a few reassessments) have been undertaken, but there has been no subsequent reassessment of the entire amphibian data set. We used version 2010.4 of the amphibian range maps, downloaded as vector polygons (shapefiles) from the IUCN Red List website (IUCN, 2010).

### Presence records

We assessed the correspondence between the amphibian range maps and two sets of presence records: a very large

data set, covering a good proportion of described amphibians but which may include inaccurate localities ('whole GBIF data'), and a smaller data set, in which we tried to minimize the number of inaccurate localities ('high-quality data').

### Whole GBIF data

The Global Biodiversity Information Facility (GBIF, <http://www.gbif.org/>) is a data portal established in 2001 to allow free and open access to biodiversity data. It is an amalgamation of non-validated organismal presence data gathered from a variety of sources. It currently holds more than 390,000,000 species distribution records across the globe; approximately half of the point localities originated from museum records and the rest from field studies (Edwards, 2004; Boitani *et al.*, 2011). The GBIF database also includes data from faunistic atlases of some European countries (e.g. Spain, Portugal). A list of current Red List amphibian taxonomy following *Amphibian Species of the World* (Frost, 2011) and primary common names was sent to GBIF, who matched the taxonomy used in the Red List with their own records and kindly provided us with all species records for which coordinates were available ( $n = 4507$  species from all continents except Antarctica).

### High-quality data

We obtained high-quality, georeferenced presence records through two main sources: the GBIF portal and the recent papers reporting accurate geographical localities (mostly from the journal *Check List*). *Check List* is a peer-reviewed, online international journal publishing lists of species and notes on the geographical distribution of any taxon. Papers are authored primarily by taxonomists; therefore, we expect a minimum rate of misidentification, although even taxonomists may have difficulties with cryptic species. We inspected all issues of the journal, from the beginning of publication (2005) to June 2011 (23 issues). Museum records digitized in GBIF are often old; older records are more likely to contain a higher proportion of positional and taxonomic errors than more recent data (Boitani *et al.*, 2011); we therefore considered only GBIF point localities collected between 1990 and 2011 for which the reported coordinate precision is 1000 m or finer. These parameters are expected to remove a large amount of inaccurate information. All data were visually inspected for any apparent discrepancies, and duplicate presence records from the same locality were omitted. We also removed all dubious records (e.g. records that probably came from sister species or from recently subdivided species) and all localities where the species is non-native (following Kraus, 2009); 0.7% of records were removed because they were dubious or non-native. Data from recent publications and GBIF were merged for analyses into a single data set ('high-quality data') which included 423 species from all continents. Ninety-three per cent of 'high-quality' data points were from GBIF, and there is therefore some overlap between the

'high-quality' and the 'whole GBIF' data. Despite some of the 'high-quality' data being included within the 'whole GBIF' data set, the two data sets represent substantially different information, as only 5% of the 'whole GBIF' data have been included in the 'high-quality' data set.

## Analyses

We assigned each species to the continent that included the centroid of the species' range, considering six continents: Europe, North and Central America (from Panama northwards), South America, Africa, Asia and Oceania (including Papua New Guinea, New Zealand and Australia). The division of continents reflected previous global analyses of amphibian expertise (Rodrigues *et al.*, 2010). We mapped species ranges using an equal-area Mollweide projection in order to calculate their areas. A spatial intersection analysis, using both amphibian ranges and the extracted point data, was carried out for every species to determine the number of records falling within the range polygon. This procedure was repeated for both the 'whole GBIF' and the 'high-quality' data sets and was done in a batch process using Python and GRASS scripts to automate tasks (see Appendix S1 in Supporting Information). The proportion of records falling within the amphibian range maps was then calculated for each species. For the 'high-quality' data set, we also calculated the Euclidean distance between each point outside the species range and the polygon's closest edge. We assigned a distance of 0 km to points within the range boundary.

The distances of 'high-quality' presence points from the boundary of the range of each species were then used to calculate the minimum distances needed to include 70%, 80%, 90% or 95% of presence records for any given species. This allowed us to identify 'confidence bands' of species ranges (considering both points inside the species range, and those outside the range within a given distance from the boundary). These bands are intended as a further measure of the degree of spatial accuracy of ranges.

Our analyses focused primarily on the extent to which range maps do not include the extant populations (omission error). We also used the  $\alpha$ -hull methodology (Burgman & Fox, 2003) to explore the extent to which range maps could over-represent these populations (i.e. commission error) – for instance, because areas within the range are actually not occupied (Appendix S2).

We used generalized linear models (GLM) to analyse factors determining the variation of range fit, considering continent and range size as independent variables. We used weighted least squares, by weighing each species depending on the number of records found. Because of overdispersion, we assumed a quasi-binomial error distribution and used an *F*-test to assess significance (Crawley, 2007). We used mixed models to analyse the variation of the distance of records from species boundaries, considering continent as fixed factor and species identity as random effect. Furthermore, all models included range area as a covariate. We then

performed Tukey's all-pair comparisons to assess the significance of differences between pairs of continents within the GLM and mixed-model frameworks (Bretz *et al.*, 2011).

We used the data published by Martin *et al.* (2012) as a measure of geographical distribution of ecological observations. These data are the geographical location of all field studies published in major ecological journals during the period 2004–2009. For each continent, we calculated the availability of ecological information, corrected for biodiversity, as the ratio of the number of ecological observation studies to the number of amphibian species. We used the ratio between the number of studies and the number of species instead of using the raw number of studies in order to take into account strong geographical differences in amphibian species richness. According to these data, Europe is the continent subject to the largest number of ecological studies, taking into account its biodiversity, followed by North America, Oceania, Asia, Africa and South America (see Martin *et al.*, 2012). We then calculated Spearman's correlation between the availability of ecological information and the measures of quality of amphibian range maps (i.e. average range fit measured using both the 'whole GBIF' and the 'high-quality' data; average distance from presence records to range edge).

We performed geographical analyses using GRASS GIS 6.4 (GRASS Development Team, 2008), ArcGIS 9.2 and ArcGIS 10.0 (ESRI, Redlands, CA, USA, 2006, 2010). We performed statistical analyses in the R 2.12 environment (R Development Core Team, 2010) using the packages LME4 (Bates & Maechler, 2010) and MULTCOMP (Hothorn *et al.*, 2008; Bretz *et al.*, 2011).

## RESULTS

### Analysis of range fit with the whole GBIF data

Overall, GBIF was the source for 206,156 presence points from 4507 amphibian species for which there are published range shapefiles in the IUCN Red List of Threatened Species (Table 1). Our data set therefore included 65% of about 7000 described amphibians (AmphibiaWeb, 2013). South America had the greatest number of species in our data set; Europe had the fewest species but the largest number of presence records. On average, 73.8% of validation records were within the species' ranges. Range fit was highest for species with large geographical ranges ( $F_{1,4500} = 2594.34$ ,  $P < 0.0001$ ). Furthermore, when taking into account differences in range size, the range fit was significantly different among continents ( $F_{5,4500} = 203.8$ ,  $P < 0.0001$ ). Europe had the highest range fit, followed by North/Central America and Oceania. Africa, Asia and South America had the lowest range fit (Table 1). Range fit was significantly correlated with measure of availability of ecological information published by Martin *et al.* (2012) ( $r_s = 0.94$ ,  $P = 0.005$ ). Range fit differed significantly between all pairs of continents, except between South America and Asia, which showed similar

**Table 1** Fit of amphibian range maps across continents.

Continent	All GBIF data			High-quality data		
	No. of species	No. of records	% within the range*	No. of species	No. of records	% within the range*
Europe	56	79,737	92.6	16	8326	94.4
North/Central America	795	59,279	75.8	116	1011	84.3
South America	1678	18,123	33.6	86	431	57.4
Africa	715	7331	44.5	90	223	77.6
Asia	839	13,296	38.1	65	674	63.7
Oceania	424	28,390	67.3	50	212	69.2

\*Weighted average across species.

proportions (post hoc comparison between Asia and South America,  $P = 0.95$ ; all other post hoc pairwise comparisons,  $P < 0.005$ ). Continent and range size explained a substantial proportion of variation ( $R^2_D = 0.60$ ).

### Analysis of range fit with high-quality data

We obtained 10,877 high-quality point localities for 423 species. North America had the largest number of species in this data set; Europe had the fewest species but the largest number of presence records (Table 1, Appendix S3). The high-quality data set showed better performance than the whole GBIF data set, with a consistently higher range fit. Nevertheless, the pattern of variation among continents was similar in the two data sets (Table 1).

Despite the fact that, for several species, a number of records were outside the species' range, validation points were generally coherent with ranges (see, for example, Fig. 1). On average, 89% of records were within the species' ranges. Range fit was positively related to the size of geographical ranges ( $F_{1,416} = 29.9$ ,  $P < 0.0001$ ). When taking into account the effect of range size, fit was significantly different among continents ( $F_{5,416} = 29.3$ ,  $P < 0.0001$ ). Europe and North America had the highest proportion of records within the range, followed by Africa and Oceania. Range fit was lowest in Asia and South America (Table 1). European, North American and Oceanian species showed similarly high range fit (post hoc, all  $P$ -values  $> 0.2$ ). European species showed higher range fit than South American, African and Asian species (post hoc, all  $P$ -values  $\leq 0.01$ ). North American and Oceanian species showed higher range fit than South American and Asian (all  $P$ -values  $< 0.03$ ), and no differences in range fit from African species ( $P \geq 0.19$ ). No significant differences were detected between Africa, Asia and South America (all  $P$ -values  $\geq 0.16$ ). Continent and range size explained 39% of variation. Range fit was significantly correlated with the availability of ecological information ( $r_s = 0.83$ ,  $P = 0.043$ ).

The number of points falling outside the range was strongly correlated with the overall number of records for that species (Pearson's correlation on log-transformed data:  $r = 0.54$ ,  $n = 423$ ,  $P < 0.0001$ ). In 55% of species, none of

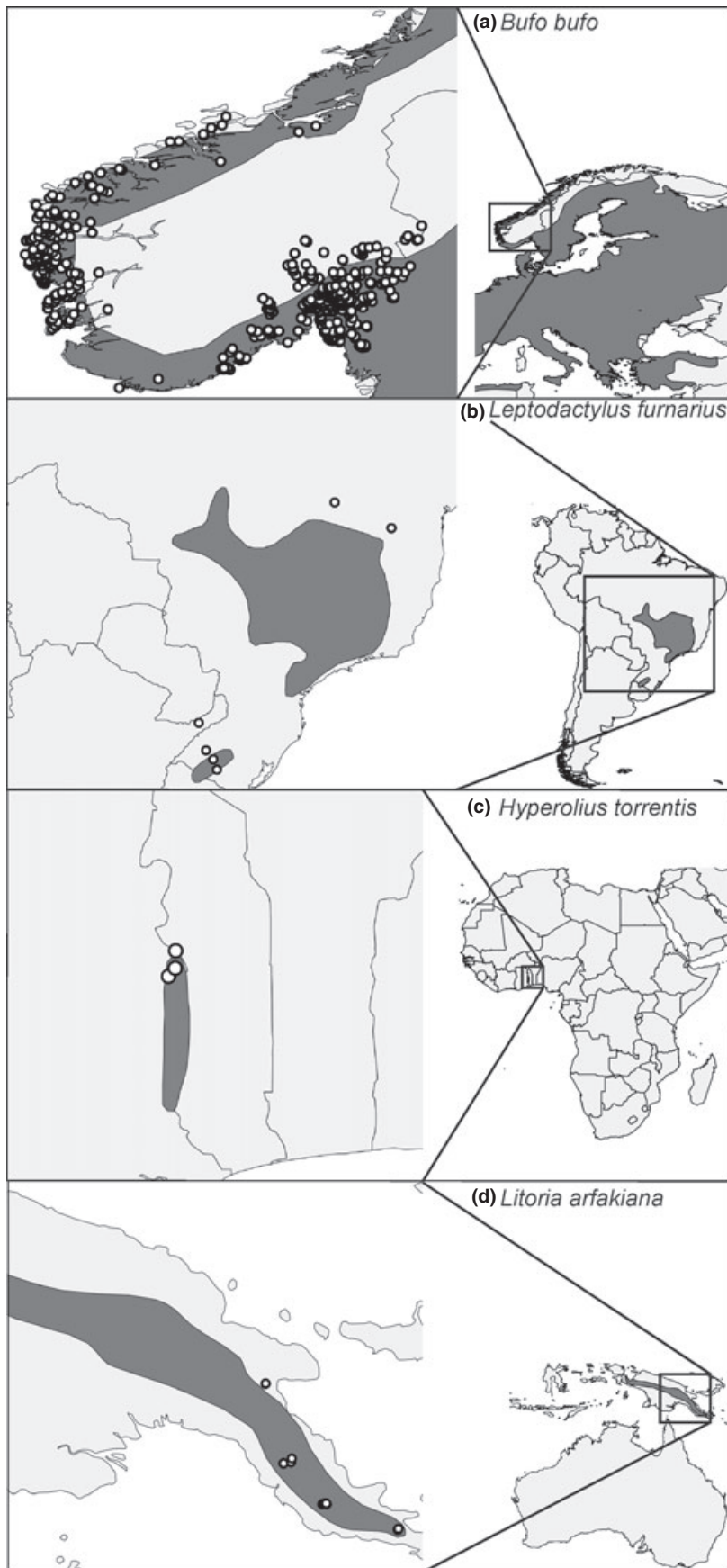
the validation points fell outside the range, but all these species had a limited number of localities (57 or fewer).

### Analysis of distances

There was a negative relationship between the average distance of presence records from the range boundary and range size ( $F_{1,416} = 47.3$ ,  $P < 0.0001$ ), indicating that the most distant points were associated with species with small ranges. Furthermore, the distance of high-quality presence records from species boundaries was significantly different among continents (mixed model,  $F_{5,416} = 11.7$ ,  $P < 0.0001$ ), South America being the continent with the highest distances (Table 2, Fig. 2). Distances to range edges were significantly higher for South American species than for species on other continents (post-hoc, all  $P$ -values  $\leq 0.001$ ). The distances of African and Asian species from range edges were higher than those of North American species (both  $P < 0.05$ ). The other differences were not significant (all  $P$ -values  $> 0.2$ ). The distances of point locality data to species' range edges were negatively correlated with the availability of ecological information ( $r_s = -0.83$ ,  $P = 0.043$ ). Overall, the average distance of high-quality records from range boundaries was strongly correlated with range fit (Spearman's correlation:  $r_s = -0.939$ ,  $n = 423$ ,  $P < 0.0001$ ), suggesting that range fit and distance from the boundary provide congruent measures of the robustness of maps.

Ninety-five per cent of all GBIF and *Check List* records were either within the range or within 31 km of the species' range border, i.e. 31 km was the confidence band including 95% of all records at the global scale. Similarly, 90% were either within the range or within 2 km (Table 2). A very small proportion (0.15%) were in the range 500–1108 km. These few very distant points were from South America (63%), Oceania (25%), Africa and Asia (6% each); all the distant points represent species in the order Anura (frogs and toads). Nevertheless, differences among continents were strong (Table 2, Fig. 2). For European species, a 2-km confidence band included 95% of points. By contrast, for South American species a 123-km band was required to include 80% of points (Table 2).

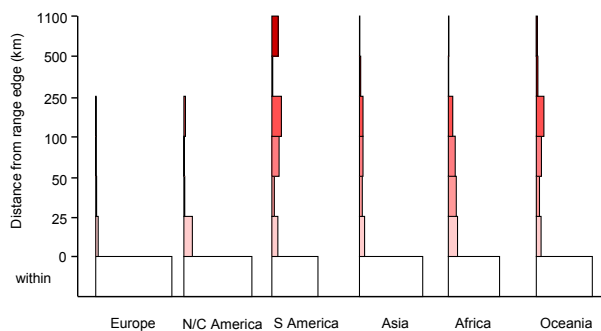




**Figure 1** Validation points for amphibian species from different continents and with ranges of different area. (a) *Bufo bufo*, Europe: range area 10,500,000 km<sup>2</sup>; (b) *Leptodactylus furnarius*, South America: range area 1,155,000 km<sup>2</sup>; (c) *Hyperolius torrentis*, Africa: range area 4700 km<sup>2</sup>; (d) *Litoria arfakiana*, Oceania: range area 199,000 km<sup>2</sup>. The scale is not the same in the four maps; darker colours represent the species range maps.

**Table 2** Interval bands representing the distance of high-quality point locality data to amphibian species range edges (in kilometres beyond the range) by continent. Percentage interval bands indicate the proportion of point records that can be found within a particular distance. Thus, for North America, 95% of all records can be found within 21 km of the edge of the species' range, while for Asia, 80% of records are within 36 km of the range edge.

Continent	Percentage interval bands (km)			
	70%	80%	90%	95%
Europe	0	0	0	2
North/Central America	0	0	6	21
South America	62	123	285	368
Africa	0	3	57	140
Asia	16	36	72	125
Oceania	3	42	141	192
All continents	0	0	2	31



**Figure 2** Frequency of high-quality point localities falling within species ranges, and outside ranges at different distance intervals from the range edge for amphibian taxa in six regions of the world. The area of each bar is proportional to the proportion of points within a given distance, as the number of validation points was different among continents. The total number of records per continent is reported in Table 1.

## DISCUSSION

Our assessment of amphibian range maps provides important information on the properties of these data. On the one hand, measures of fit suggest that range maps represent the actual distribution of most amphibians well, and that omission errors are relatively limited, particularly in certain geographical areas. Even for the few species in which several records were rather far from the range boundary (e.g. *Leptodactylus furnarius*, Fig. 1b), the range still provides a reasonable and useful picture of species distribution. On the other hand, the fit between range maps and distribution suggests strong variation among continents, and geographical variation parallels observed global bias for research efforts and ecological observations (Martin *et al.*, 2012). The current study provides measures of range accuracy that may be useful for future research using the amphibian range maps as

baseline data, and demonstrates a need for greater investment in the continuous update and improvement of maps, particularly in the megadiverse areas of tropical Asia and South America.

Using the 'whole GBIF data' provided poorer measures of fit when compared with the 'high-quality data' that only includes recent and more accurate data points. Using GBIF as an unscreened source of information can be problematic, as data can be provided by multiple contributors and no filtering or error-checking procedure is implemented; therefore, a significant proportion of points may be affected by spatial errors (inaccurate coordinates) (Yesson *et al.*, 2007). Accuracy issues may be particularly problematic in old records, because of both positional errors and taxonomic changes (Boitani *et al.*, 2011). Using all GBIF data yielded a very large validation data set, with > 200,000 records covering the majority of amphibian species, but the inclusion of unverified records limited the usefulness of these data to assess amphibian range maps, and might have increased bias in the measure of range fit. For these reasons, GBIF records cannot be used uncritically to update species ranges (e.g. in procedures of automated update of point locality data), because errors exist in the database. The range fit was much higher in the 'high-quality data', which was obtained in part from the GBIF, but in this instance data from GBIF underwent some screening. The use of selected, recent data reduced sample sizes and required substantial operator workload, but limited errors, confirming the importance of data quality in spatial analyses validating the performance of any kind of distribution models (including range maps) (Boitani *et al.*, 2011). These records, and particularly those from specific publications (e.g. the journal *Check List*), might be used to improve or refine ranges.

The development of the range maps involved several steps that could account for some of the errors observed across all continents (Chanson *et al.*, 2008b; IUCN Standards & Petitions Subcommittee, 2011). First, given the methodology used to produce range maps (workshops), and given issues of scale with hand-drawn boundaries (and transferring them to a digital format), a degree of error was to be expected in the maps (Sanderson *et al.*, 2002). Second, when the maps were originally drawn, access to GIS tools was not as easy as it is today (Chanson *et al.*, 2008b). Third, it is possible that experts in developing countries such as those in South America and Asia may not have had access to many primary sources of records or primary literature at the time that the maps were compiled. Taxonomic uncertainty may be an additional cause of spatial error, as it may result in modifications and splitting of ranges on the basis of expert opinion. Furthermore, the accuracy of ranges reflects the global bias towards better-sampled regions of the world with easier access to information (Martin *et al.*, 2012), and possibly certain taxonomic groups as well. For instance, the continent with highest range fit (i.e. Europe) is also the continent with most observations (Table 1; see discussion below). An additional source of variation is the procedure of map-drawing.

In North America, in some instances, ranges were expanded to match administrative boundaries, possibly resulting in inflated ranges and reduced omission error (i.e. when a species is incorrectly thought to be absent) arguably at the cost of an increase of commission errors (i.e. when a species is incorrectly thought to be present) (Boitani *et al.*, 2011). Mapping strategies that use fine cartographical grids may provide more effective results.

The quantification of error here focused on range limits. Unused areas inside ranges (due to unsuitable habitat or other limiting factors) are not accounted for. While these may be acceptable for large-scale biogeographical analyses, they are known to affect conservation prioritization analyses (Hurlbert & White, 2005; Rondinini *et al.*, 2005, 2011; Hurlbert & Jetz, 2007; Cantú-Salazar & Gaston, 2013). To overcome these issues, amphibian ranges may be combined with habitat suitability models that integrate information on factors such as habitat availability and climatic suitability, where available. Owing to the scarcity of information, broad-scale deductive models may be built by overlaying areas of suitable or unsuitable habitat (e.g. based on water availability, intensely urbanized areas and, if available, habitat maps) to range polygons, to provide some preliminary indication of possible species occupancy within a polygon (Rondinini *et al.*, 2006, 2011). This would provide a more robust estimate of actual species distribution, identifying unsuitable areas within amphibian ranges and thus refining the information available for conservation. Unfortunately, data on the micro-habitat preferences, tolerance of habitat disturbance, and climatic envelope are extremely scarce for many amphibians, thereby limiting the possibility of using them to improve range maps.

We attempted to assess commission errors within ranges using  $\alpha$ -hulls created on the basis of presence points. This approach may allow unoccupied areas within a species' range to be excluded (Burgman & Fox, 2003), and may provide more refined information on patterns of occupancy within the species ranges (Appendix S2). The  $\alpha$ -hull analysis suggests that range maps might overestimate species distribution, and overestimation might be particularly relevant for species with large geographical ranges (Appendix S2). However, the results of the  $\alpha$ -hull analysis should be considered with caution:  $\alpha$ -hulls are an approximate measure of species distribution, and simulations suggested they can have good performance when presence points are numerous (20,000 or more per species) and are distributed across the species' whole range (Burgman & Fox, 2003). Unfortunately, such a large quantity of high-quality presence data was not available for any species. Additional analyses, based on fine-scale gridded surveys (e.g. from high-resolution atlases) covering multiple continents, are needed for a more comprehensive assessment of unoccupied areas within ranges (Hurlbert & White, 2005; Hurlbert & Jetz, 2007).

Overall, the amphibian ranges adequately represent the limits of species distributions considering our current taxonomic knowledge of these species (see below for further

discussion on the implications of taxonomic resolution), but the results of this analysis can be used to improve their use. To reduce the risk of omitting amphibian populations from analyses, researchers may integrate ranges with confidence bands. Bands might be continent-specific, ranging from a few kilometres in Europe and North America, to 50–100 km in South America and Oceania (Table 2). Clearly, the decision of confidence bands to be used (e.g. 95%, 80% bands) would depend on the aims of the analyses, and on the relative cost of omission versus commission errors (Loiselle *et al.*, 2003; Rondinini *et al.*, 2006). When commission errors are more dangerous than omissions, it may be sensible to use ranges as they are, provided it is clear that they represent a lower limit of species distributions.

As expected, the geographical distribution of number of records and range fit follows the global pattern in knowledge and (inversely) in species richness (Table 1) (Rodrigues *et al.*, 2010; Martin *et al.*, 2012): range fit was highest in Europe and North America, where most amphibian specialists are based, where more field studies are performed (Rodrigues *et al.*, 2010; Martin *et al.*, 2012), and thus where amphibian distributions are best known. Available data on the distribution of ecological observations (Martin *et al.*, 2012) are certainly not complete, and are likely to omit some biodiversity surveys and monitoring efforts occurring in tropical countries (Young *et al.*, 2001); nevertheless, these data are probably a good representation of the global bias in biodiversity knowledge. This is supported by the similarity between the global patterns of ecological observations and taxonomic expertise (cf. Rodrigues *et al.*, 2010; Martin *et al.*, 2012). Ecogeographical factors may also be implicated in the differences in range fit among continents. Many amphibian ranges tend to be smaller in tropical regions, where it is possible to find more micro-endemics (Whitton *et al.*, 2012), and range fit was lower for small ranges. This may contribute to the limited range fit in the tropical continents. South America and Asia were the continents with lowest range fit (Tables 1 & 2). This can be partly explained by the large number of very recent data on amphibian distributions in South America, which have improved and expanded the available distribution data. For instance, the journal *Check List* is based in Brazil, and a large proportion of papers focus on Latin America; an increasing number of studies and records from South America have also been published in other journals covering similar topics, such as *Herpetological Review*. A similar situation probably occurs in Asia, which includes both rich countries receiving substantial research efforts and tropical developing countries. In general, and taking into account continental biodiversity, South America is the continent where the fewest field studies have been undertaken and/or made public, while Asia and Africa are the continents with fewest studies considering their area (Martin *et al.*, 2012). An additional potential reason for the relatively low range fit in South America, Asia and Africa is a more limited access to data sources (museum records, primary literature) for the initial range maps. It is also possible



that in these continents, our measures of range fit might be slightly overestimated because some of the observation points used for validation were also available to the experts that drew the maps, although we would expect a degree of variation across regions, taxa and species with differing range sizes, given that range revisions were conducted at assessment workshops with other experts who contributed with additional information. With the tropical areas of these three continents being hotspots for amphibian species richness, and with there being many unexplored areas with the potential to harbour more new records as well as new species, the funding available for researchers to investigate areas that remain poorly studied remains insufficient.

Furthermore, and particularly in tropical regions, incomplete taxonomic delimitations could seriously affect amphibian conservation (Vieites *et al.*, 2009; Angulo & Icochea, 2010). The lack of taxonomic resolution of unrecognized cryptic taxa within widely distributed nominal species (complexes), for instance, impedes the assessment of the conservation status of the cryptic taxa under these names, and underestimates their extinction risks. The conservation status of groups known to be taxonomically complex but that have been assessed as of Least Concern should be treated with caution (Angulo & Icochea, 2010), particularly if studies on intraspecific diversity are lacking. Taxonomic knowledge is strongly related to the accuracy of range maps: as taxonomic uncertainty can overestimate distribution, it can cause errors in the assessment of the appropriate category of threat, with a significant impact on conservation decisions (Angulo & Icochea, 2010). Refining taxonomic resolution is not an easy task for many species complexes (Angulo & Icochea, 2010). Greater collaborations among taxonomists and regional biologists, as well as joint efforts between taxonomists and biogeographers are thus needed to improve our knowledge of species identity and distribution of amphibians, in order to better inform our assessments of extinction risk and the resulting conservation decisions.

This study identifies regions where it would be useful to invest resources into improving range accuracy and reassessing species, and stresses the importance of keeping ranges up to date. While this may appear relatively straightforward, it requires significant additional resources, and resources for amphibian assessments have become increasingly difficult to secure in the context of the current economic climate and also because the group has already been assessed at least once. Under these circumstances, alternative strategies for maintaining the global amphibian database and maps need to be explored, and the identification of well-organized and coordinated regional and national expert groups that would be willing to take stewardship of their assessments is a first step in this direction. However, additional resources are still needed to support these groups and to support both assessment and map updates. Our analysis focused on amphibian ranges; range maps of other taxa, such as mammals, birds and a number of reptiles, are made available by the IUCN

and are used in a wide range of studies, and could also be analysed in a similar fashion to that applied here.

Uncertainty is an inherent component of biological studies, and also affects amphibian ranges. However, not all uncertainty is equal: for example, Akçakaya *et al.* (2000) identified at least three different types of uncertainty associated with conservation assessments (semantic uncertainty, measurement error and natural variability). For amphibian ranges, much of the uncertainty can be associated with aspects that can be aggregated under measurement error (e.g. the various stages and approaches involved with drawing polygons), but also natural variability (coming with soft range boundaries that are in flux). Understanding the nature of the uncertainty involved may help interpret existing results and direct future biodiversity research.

## ACKNOWLEDGEMENTS

We thank three referees for comments on early versions of the manuscript. N. Cox and S. Stuart provided useful information and suggestions on early phases of the study; A. Corti helped with data collection; and A. Hughes helped with GBIF data analyses. G.F.F. was funded by a scholarship of the University of Milano-Bicocca.

## REFERENCES

- Akçakaya, H.R., Ferson, S., Burgman, M.A., Keith, D.A., Mace, G.M. & Todd, C.R. (2000) Making consistent IUCN classifications under uncertainty. *Conservation Biology*, **14**, 1001–1013.
- AmphibiaWeb (2013) *AmphibiaWeb: information on amphibian biology and conservation*. University of California, Berkeley, CA. Available at: <http://amphibiaweb.org/>.
- Angulo, A. & Icochea, J. (2010) Cryptic species complexes, widespread species and conservation: lessons from Amazonian frogs of the *Leptodactylus marmoratus* group (Anura: Leptodactylidae). *Systematics and Biodiversity*, **8**, 357–370.
- Bates, D. & Maechler, M. (2010) *lme4: linear mixed-effects models using S4 classes*. R package version 0.999375-37. Available at: <http://cran.r-project.org/web/packages/lme4/>.
- Becker, C.G. & Zamudio, K.R. (2011) Tropical amphibian populations experience higher disease risk in natural habitats. *Proceedings of the National Academy of Sciences USA*, **108**, 9893–9898.
- Boitani, L., Maiorano, L., Baisero, D., Falcucci, A., Visconti, P. & Rondinini, C. (2011) What spatial data do we need to develop global mammal conservation strategies? *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2623–2632.
- Bretz, F., Hothorn, T. & Westfall, P. (2011) *Multiple comparisons in R*. Taylor & Francis, Boca Raton, FL.
- Burgman, M.A. & Fox, C.W. (2003) Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. *Animal Conservation*, **6**, 19–28.

- Cantú-Salazar, L. & Gaston, K.J. (2013) Species richness and representation in protected areas of the Western hemisphere: discrepancies between checklists and range maps. *Diversity and Distributions*, **19**, 782–793.
- Chanson, J.S., Hoffman, M., Cox, N. & Stuart, S.N. (2008a) The state of the world's amphibians. *Threatened amphibians of the world* (ed. by S.N. Stuart, M. Hoffman, J.S. Chanson, N.A. Cox, R.J. Berridge, P. Ramani and B.E. Young), pp. 33–44. Lynx Editions, IUCN and Conservation International, Barcelona.
- Chanson, J.S., Stuart, S.N., Cox, N., Young, B.E. & Hoffman, M. (2008b) The Global Amphibian Assessment (GAA): history, objectives and methodology. *Threatened amphibians of the world* (ed. by S.N. Stuart, M. Hoffman, J.S. Chanson, N.A. Cox, R.J. Berridge, P. Ramani and B.E. Young), pp. 30–32. Lynx Editions, IUCN and Conservation International, Barcelona.
- Cooper, N., Bielby, J., Thomas, G.H. & Purvis, A. (2008) Macroecology and extinction risk correlates of frogs. *Global Ecology and Biogeography*, **17**, 211–221.
- Crawley, M.J. (2007) *The R book*. Wiley, Chichester, UK.
- Edwards, J.L. (2004) Research and societal benefits of the Global Biodiversity Information Facility. *BioScience*, **54**, 485–486.
- Ficetola, G.F., Bonardi, A., Sindaco, R. & Padoa-Schioppa, E. (2013) Estimating patterns of reptile biodiversity in remote regions. *Journal of Biogeography*, **40**, 1202–1211.
- Finer, M., Jenkins, C.N., Pimm, S.L., Keane, B. & Ross, C. (2008) Oil and gas projects in the western Amazon: threats to wilderness, biodiversity, and indigenous peoples. *PLoS ONE*, **3**, e2932.
- Foody, G.M. (2011) Impacts of imperfect reference data on the apparent accuracy of species presence–absence models and their predictions. *Global Ecology and Biogeography*, **20**, 498–508.
- Frost, D.R. (2004) *Amphibian species of the world: an online reference*. Version 3 (22 August, 2004). American Museum of Natural History, New York. Available at: <http://research.amnh.org/vz/herpetology/amphibia/>.
- Frost, D.R. (2011) *Amphibian species of the world: an online reference*. Version 5.5 (31 January, 2011). American Museum of Natural History, New York. Available at: <http://research.amnh.org/vz/herpetology/amphibia/>.
- Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.
- GRASS Development Team (2008) *Geographic Resources Analysis Support System (GRASS) Software*. Open Source Geospatial Foundation Project, Beaverton, OR. Available at: <http://grass.osgeo.org/>.
- Grenyer, R., Orme, C.D.L., Jackson, S.F., Thomas, G.H., Davies, R.G., Davies, T.J., Jones, K.E., Olson, V.A., Ridgely, R.S., Rasmussen, P.C., Ding, T.-S., Bennett, P.M., Blackburn, T.M., Gaston, K.J., Gittleman, J.L. & Owens, I.P.F. (2006) Global distribution and conservation of rare and threatened vertebrates. *Nature*, **444**, 93–96.
- Hof, C., Araújo, M.B., Jetz, W. & Rahbek, C. (2012) Additive threats from pathogens, climate and land-use change for global amphibian diversity. *Nature*, **480**, 516–519.
- Hothorn, T., Bretz, F. & Westfall, P. (2008) Simultaneous inference in general parametric models. *Biometrical Journal*, **50**, 346–363.
- Hurlbert, A.H. & Jetz, W. (2007) Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Sciences USA*, **104**, 13384–13389.
- Hurlbert, A.H. & White, E.P. (2005) Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. *Ecology Letters*, **8**, 319–327.
- IUCN (2010) *IUCN Red List of Threatened Species 2010.2*. IUCN, Gland, Switzerland. Available at: <http://www.iucnredlist.org/>.
- IUCN Standards and Petitions Subcommittee (2011) *Guidelines for using the IUCN Red List categories and criteria*. Version 9.0. IUCN, Gland, Switzerland. Available at: <http://www.iucnredlist.org/documents/RedListGuidelines.pdf>.
- Kanda, L.L., Fuller, T.K., Sievert, P.R. & Kellogg, R.L. (2009) Seasonal source–sink dynamics at the edge of a species' range. *Ecology*, **90**, 1574–1585.
- Kraus, F. (2009) *Alien reptiles and amphibians: a scientific compendium and analysis*. Springer, Dordrecht.
- Lawler, J.J., Shafer, S.L., Bancroft, B.A. & Blaustein, A.R. (2010) Projected climate impacts for the amphibians of the western hemisphere. *Conservation Biology*, **24**, 38–50.
- Loiselle, B.A., Howell, C.A., Graham, C.H., Goerck, J.M., Brooks, T., Smith, K.G. & Williams, P.H. (2003) Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology*, **17**, 1591–1600.
- Lomolino, M.V. (2004) Conservation biogeography. *Frontiers of biogeography: new directions in the geography of nature* (ed. by M.V. Lomolino and L.R. Heaney), pp. 293–296. Sinauer, Sunderland, MA.
- Martin, L.J., Blossey, B. & Ellis, E. (2012) Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. *Frontiers in Ecology and the Environment*, **10**, 195–201.
- Mokany, K. & Ferrier, S. (2011) Predicting impacts of climate change on biodiversity: a role for semi-mechanistic community-level modelling. *Diversity and Distributions*, **17**, 374–380.
- Muths, E., Gallant, A.L., Campbell, G.E., Battaglin, W.A., Green, D.E., Staiger, J.S., Walls, S.C., Gunzburger, M.S. & Kearney, R.F. (2006) *The Amphibian Research and Monitoring Initiative (ARMI): 5-year report*. Scientific Investigations Report 2006–5224. U.S. Geological Survey, Reston, VA.
- Naka, L.N., Bechtoldt, C.L., Henriques, L.M.P. & Brumfield, R.T. (2012) The role of physical barriers in the location of avian suture zones in the Guiana Shield, Northern Amazonia. *The American Naturalist*, **179**, E115–E132.

- R Development Core Team (2010) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.
- Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jiménez-Valverde, A., Ricotta, C., Bacaro, G. & Chiarucci, A. (2011) Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Progress in Physical Geography*, **35**, 211–226.
- Rodrigues, A.S.L., Gray, C.L., Crowter, B.J., Ewers, R.M., Stuart, S.N., Whitten, T. & Manica, A. (2010) A global assessment of amphibian taxonomic effort and expertise. *BioScience*, **60**, 798–806.
- Rondinini, C., Stuart, S. & Boitani, L. (2005) Habitat suitability models and the shortfall in conservation planning for African vertebrates. *Conservation Biology*, **19**, 1488–1497.
- Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H. & Possingham, H.P. (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters*, **9**, 1136–1145.
- Rondinini, C., Di Marco, M., Chiozza, F., Santulli, G., Baisero, D., Visconti, P., Hoffman, M., Schipper, J., Stuart, S.N., Tognelli, M.F., Amori, G., Falcucci, A., Maiorano, L. & Boitani, L. (2011) Global habitat suitability models of terrestrial mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2633–2641.
- Sandel, B., Arge, L., Dalsgaard, B., Davies, R.G., Gaston, K.J., Sutherland, W.J. & Svenning, J.-C. (2011) The influence of Late Quaternary climate-change velocity on species endemism. *Science*, **334**, 660–664.
- Sanderson, E.W., Redford, K.H., Chetkiewicz, C.L.B., Medellin, R.A., Rabinowitz, A.R., Robinson, J.G. & Taber, A.B. (2002) Planning to save a species: the jaguar as a model. *Conservation Biology*, **16**, 58–72.
- Stuart, S.N., Chanson, J.S., Cox, N.A., Young, B.E., Rodriguez, A.S.L., Fischman, D.L. & Waller, R.W. (2004) Status and trends of amphibian declines and extinctions worldwide. *Science*, **306**, 1783–1786.
- Stuart, S.N., Hoffman, M., Chanson, J.S., Cox, N.A., Berridge, R.J., Ramani, P. & Young, B.E. (eds) (2008) *Threatened amphibians of the world*. Lynx Editions, IUCN and Conservation International, Barcelona.
- Thomas, C.D. (2010) Climate, climate change and range boundaries. *Diversity and Distributions*, **16**, 488–495.
- Vieites, D.R., Wollenberg, K.C., Andreone, F., Köhler, J., Glaw, F. & Vences, M. (2009) Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proceedings of the National Academy of Sciences USA*, **106**, 8267–8272.
- Whitton, F.J.S., Purvis, A., Orme, C.D.L. & Olalla-Tárraga, M.Á. (2012) Understanding global patterns in amphibian geographic range size: does Rapoport rule? *Global Ecology and Biogeography*, **21**, 179–190.
- van Wilgen, N.J., Roura-Pascual, N. & Richardson, D.M. (2009) A quantitative climate-match score for risk-assessment screening of reptile and amphibian introductions. *Environmental Management*, **44**, 590–607.
- Yesson, C., Brewer, P.W., Sutton, T., Caithness, N., Pahwa, J.S., Burgess, M., Gray, W.A., White, R.J., Jones, A.C., Bisby, F.A. & Culham, A. (2007) How global is the Global Biodiversity Information Facility? *PLoS ONE*, **2**, e1124.
- Young, B.E., Lips, K.R., Reaser, J.K., Ibáñez, R., Salas, A.W., Cedeño, J.R., Coloma, L.A., Ron, S., La Marca, E., Meyer, J.R., Muñoz, A., Bolaños, F., Chaves, G. & Romo, D. (2001) Population declines and priorities for amphibian conservation in Latin America. *Conservation Biology*, **15**, 1213–1223.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** GRASS script to extract geographical information from point presence localities.

**Appendix S2** Use of  $\alpha$ -hulls to estimate the extent to which range maps over-represent distribution.

**Appendix S3** Summary of high-quality data used for analyses.

## BIOSKETCH

**G. Francesco Ficetola** is a post-doctoral researcher combining multiple approaches (landscape ecology, macroecology and evolutionary ecology) to improve knowledge of the ecology and conservation of amphibians and reptiles.

Author contributions: G.F.F., C.R., V.K., E.P.-S. and A.A. designed the study; G.F.F., A.B. and V.K. performed analyses; G.F.F. wrote the first draft of the paper; all authors discussed the results, commented on earlier versions of the manuscript, and have read and approved the final version of the manuscript.

Editor: Richard Pearson