ILORA: A database of alien vascular flora of India

Vidushi Pant[1¶], Chinmay Patwardhan[2,#], Kshitij Patil[2], Amiya Ranjan Bhowmick[2], Abhishek Mukherjee[3], Achyut Kumar Banerjee[4¶*]

[1] Sustainable India Finance Facility, Guntur, Andhra Pradesh, Pin – 522034, India

[2] Department of Mathematics, Institute of Chemical Technology, Mumbai, Maharashtra, Pin – 400019, India

[3] Agricultural and Ecological Research Unit, Indian Statistical Institute, Giridih, Jharkhand, Pin – 815301, India

[4] School of Life Sciences, Sun Yat-sen University, 135 Xingangxi Road, City - Guangzhou, Province - Guangdong, Pin – 510275, China

[#]Current address: Department of Mathematics, Karlsruhe Institute of Technology, Karlsruhe, Pin – 76131, Germany

* Corresponding author

E-mail: banerjeeachyut31@outlook.com (AKB)

[¶]These authors contributed equally to this work.

Abstract

Plant invasions pose an unprecedented threat to biodiversity and ecosystems at different spatial scales, especially for a biodiversity-rich developing nation like India. While country-level checklists are considered as the first step for effective management of plant invasions, databases having biological and ecological attributes of these plant species are of paramount importance for facilitating research activities and developing policy interventions concerning alien plant invasion in a country or region. Though checklists with the names and basic details of alien plant species are available for India, a database aggregating attributes relevant to the understanding of the invasion continuum is lacking.

We have curated data for 13 variables related to ecology, biogeography, introduction history, socio-economy and environmental niche of 1747 alien vascular plant species from 22 national and global sources to produce the Indian Alien Flora Information (ILORA) version 1.0 database. This paper describes the detailed methodology of curating these data along with the logic behind selecting these variables to be included in the database. The database, first of its kind for the Indian alien flora, will provide easy access to high quality data, offer a ready reference to comprehend the existing scenario of alien plant species in the country and provide opportunities to identify, scrutinize and prioritize the management of these species. The database is dynamic and will be continuously updated with the availability of new data. It has the provision for bidirectional information exchange with the users, which will allow increasing the resolution of the database as well as the expansion of its capacity. All data will be associated with the source information and ILORA will encourage users to refer to the primary sources to give the contributors full credit and acknowledgement. The database is, therefore, envisaged to become a nationwide collaborative platform for a wide spectrum of stakeholders involving the scientific community, citizens, and policymakers. ILORA is freely accessible via an online data repository as well as through a dedicated website (https://ilora2020.wixsite.com/ilora2020).

Introduction

India harbours 11% of the world's plant diversity and part of four global biodiversity hotspots [1]. The history of biological invasion and the subsequent colonization of the country dates back nearly eight centuries, starting with the Arabs in around 1200 followed by the Turks, Mughals, and the Europeans (https://www.britannica.com/place/India/History, accessed on 10 February 2020). Numerous plant species were introduced into India during this time. Botanical imperialism [2] reached its zenith during the British rule with the establishment of networks of botanical institutions facilitating the introduction of alien plant species with commercial or medicinal values. Some of these introduced species eventually escaped cultivation to become problematic invaders in the country. Famous examples include *Lantana camara*, which was introduced for ornamental purposes from South America (via Europe) in the early 19th century, and by the end of the same century, it was recognized as an invasive plant in India [3].

Introduction of alien plant species and the subsequent naturalization of a subset of these species remain equally, if not more, pertinent in the modern times. Biological invasion has been described as a defining feature of the Anthropocene [4]. In fact, 37% of the globally recorded alien species have been introduced in the last few decades [5]. Recent studies suggest that emerging economies, like India, are facing the greatest threat of plant invasion as a consequence of increasing globalization of trade as well as the changing climate [6, 7]. This necessitates the creation of a country-level repository of alien flora as the first step towards identifying emerging invaders and managing the ones that have already established [8]. Although scientific documentation of Indian alien flora started in the mid-20th century [9, 10], the Alien Flora of India [11], published in 2012, was the first country-wide checklist of alien vascular plant species (which also categorized the species along the introduction-naturalization-invasion continuum). This was followed by the recently published Indian version of the Global Register of Invasive and Introduced Species (hereafter, GRIIS) [12]. However, the biological and ecological attributes of these alien plant species remain scattered across a multitude of databases rendering their access tedious. Some of this information, although retrievable from global databases like CABI-Invasive Species Compendium (hereafter, CABI-ISC: www.cabi.org/isc) and Global Invasive Species Database (GISD: http://www.iucngisd.org/gisd/), often lack resolution at the national or regional levels, particularly for India (but see The Global Naturalized Alien Flora database, hereafter GloNAF; [13]). For example, the Global Alien Species First Record Database contains introduction time records for only 87 vascular plants of India [14]. Further, data extraction often involves intensive and informed query-directed search process to find specific information for a subset of species. Easy access to such information at a finer scale is of paramount importance for developing country-level policy interventions and facilitating research and development concerning the alien species of a country or region. Recognizing these necessities of scale and speed of information retrieval, comprehensive databases are being developed focusing on national or regional levels [15, 16], but see [17]; however, similar accounts do not exist for India.

Here, we introduce the **I**ndian A**l**ien Fl**o**ra Infor**m**a**t**ion (ILORA) Version 1.0 Database, a dynamic and open-source repository of comprehensive information on taxonomy, general biology, introduction pathway, economic usage, geographic distribution, and realized climatic niche of 1747 alien vascular plant species of India. The detailed methodology of curating the data and the logic behind choosing the variables have been explained to provide transparency as well as repeatability of the process. ILORA should not be considered as another checklist of alien flora of India, rather we envisage that a broad spectrum of users will find ILORA as an accessible tool to extract information about the alien plant species present in the country and explore its potential to become a collaborative

platform for the scientific community, citizens, and stakeholders for further improvement of the data quality. ILORA is freely accessible through the online repository figshare (https://figshare.com/s/ce07ef7a9a19768d8f5e; DOI: 10.6084/m9.figshare.13677460) as well as through a dedicated website; the latter will be updated regularly with contributions from both authors and future users.

Materials and Methods

Categorization of the alien flora of India

We developed a multistep pipeline for categorization of alien plant species of India (Fig 1). In the first step, three nationwide checklists, viz. Alien flora of India (1599 species; [11]), Naturalized alien flora of the Indian states (499 species; [17]), and the GRIIS- India Version 1.3 (2082 species; [12]), were consulted to create a comprehensive list of alien plant species of India (n=4180). To remove ambiguity and orthographical errors in the plant nomenclature, automated standardization of taxonomic names was conducted using the *WorldFlora* package [18] in R version 4.0.2 [19] (hereafter, R). This package validates the species names against a static version of the World Flora Online (WFO) taxonomic backbone data, which is actively curated by global experts based on The PlantList backbone (http://www.theplantlist.org/) (static since 2013), and therefore provides the most updated and comprehensive taxonomic reference of vascular plants [20]. The list was further cleaned by removing duplicates and synonyms, and the infraspecific taxa and artificially hybridized species were not considered further (n=1677).

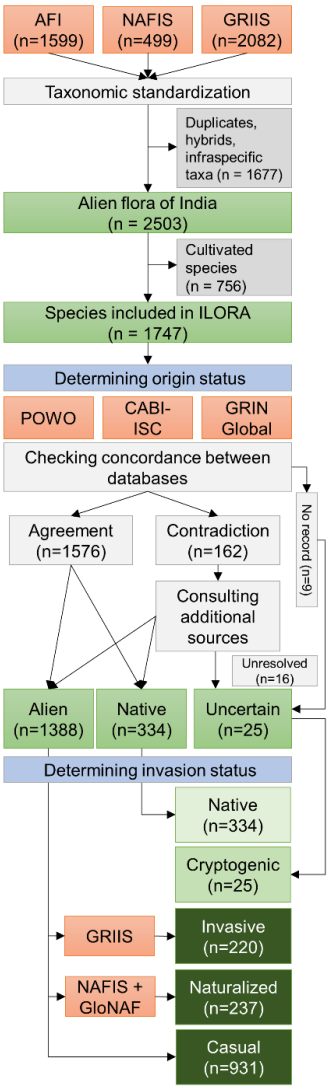

Fig 1. The methodology followed for the categorization of the Indian alien flora in the ILORA database. The method involved four steps: species assembling (compiling the alien vascular flora of India), taxonomic standardization (standardizing the accepted scientific names of the species), nativity classification (classifying the species based on their origin status) and finally, species categorization (categorizing the species based on their invasion status in India).

The selected plants (n=2503) were subjected to a two-step verification process (Fig.1) involving national-level checklists and global databases to determine their – i) origin status (i.e., is the species native or alien to India) and ii) invasion status (i.e., what is the degree of naturalization of the species) [21] in India. To designate the species categories, we followed the Darwin Core (dwc) standard and adopted the proposed controlled vocabulary for dwc:establishmentMeans [22] to maintain stronger links to existing terms and ensure future interoperability across multiple databases. First, the cultivated aliens were identified from the Alien Flora of India (n=777). Among these, the species which were identified as naturalized and/or invasive in either of the four databases, namely the Naturalized alien flora of the Indian states databases, GloNAF, GRIIS and the Germplasm Resource Information Network (hereafter, GRIN: https://www.ars-grin.gov; accessed on 4 April 2020), were removed from the list of

the cultivated species (n=21). The cultivated aliens thus identified (n=756) were retained in this version, however, no variable information was extracted for these species.

The origin status of the remaining species (n=1747) was ascertained using three global databases, namely, POWO, CABI-ISC and GRIN (accessed on 20 April 2020). These databases were chosen because they maintain detailed profiles of a large number of plant species which are frequently updated, provide traceable data sources, and are widely referred by plant biologists to trace the origin of a species [23, 24]. If the origin status of a species was identified as 'native' or 'alien' to India in at least one of these databases and was not contradicted by the remaining (n=1576), it was accepted. However, in case of any contradiction between databases (e.g., two databases identified a species as alien whereas one identified it as native) (n=162), peer-reviewed journal articles, published books, and online databases were consulted to establish the origin status of the species. For literature and book search, the Google Scholar database (https://scholar.google.com/; accessed on 29 April 2020) was used. The search terms were optimized following a preliminary search based on a handful of relevant articles retrieved by using the names of the species. The optimized search string consists of a combination of species name and nativity (and its synonyms and related words). The advanced search function of the Google Scholar database was used to search for “species name” with at least one of the words “native” OR “endemic” OR “native region” OR “nativity” OR “origin” present anywhere in the article. The search results were screened until the origin status of the species was discerned confidently. The online databases and repositories were retrieved through a Google search (http://www.google.com/; accessed on 29 April 2020) and the databases were considered based on the same criteria (i.e., frequently updated and with traceable data sources). The species with ambiguous origin status (i.e., which cannot be resolved by the global databases and literature search) were categorized as 'cryptogenic' (i.e., uncertain biogeographic status) following [25]. At the end of this step, the origin status of the alien species was categorized as – alien (n=1388), native (n=334) and cryptogenic (n=25).

In the second step, the status of the alien species in the introduction-naturalization-invasion continuum was determined (Fig.1). From the list of alien species (n=1388), the 'invasive' species were categorized by consulting the GRIIS database (i.e., the species identified as ‘Alien Invasive’ in GRIIS) (n=220) as it provides the latest list of invasive alien species present in India. The species which have been identified as naturalized aliens by the Naturalized Alien Flora checklist and the GloNAF database, were categorized as ‘naturalized’ (involving both invasive and non-invasive naturalized ones) (n=237). The remaining alien species were identified as 'casual alien' following [26] (n=931). Finally, at the end of this step, the assembled list of alien flora of India was classified into six categories, namely invasive, naturalized, casual alien, cryptogenic, native, and cultivated species. We included the native species in this database version because these species were identified as aliens in one of the three national checklists and expert in-depth knowledge in future may provide the conclusive status of these species (see ‘Future directions’ section).

Data curation

Taxonomic classification and species information

Taxonomic classification is essential for identification and analysis of evolutionary relationships between the species and hence, is an indispensable part of any plant-based database. The taxonomic information (class and order) was recorded from POWO (and from CABI-ISC, if not available in POWO) and the genus, species and family names were retained as obtained from the standardization of species names using the WFO taxonomic backbone. The phylogenetic relationship

between the three categories of alien species (invasive, naturalized and casual) was resolved through 'mega-tree' approach by using the *V.PhyloMaker* package [27] in R (specifically, using the options of build.nodes.1 and scenario 3). The common names of the species (in English) were documented from GRIN, POWO and CABI-ISC databases. Since, the vernacular names are easier to relate to and may increase the accessibility and usability of the database, we curated the vernacular names of the species from two online repositories, namely the eFlora of India (https://efloraofindia.com/; accessed on 31 May 2020) and Flowers of India (http://flowersofindia.net/; accessed on 31 May 2020), and the digital database of Flora of Peninsular India (http://flora-peninsula-indica.ces.iisc.ac.in/welcome.php; accessed on 31 May 2020). The vernacular names were standardized using the ISO 639-3:2007 code consisting of three-letter language identifiers for the representation of languages.

Along with taxonomic identity, the general information of each species, namely its growth habit, duration (perennial, annual, biennial) and group (monocot/dicot), was collated from the TRY Database [28] (accessed on 23 May 2020), a comprehensive web-archive that integrates nearly 400 different plant traits. If there was missing data in TRY, databases focused on Indian flora, namely the digital databases of Flora of Peninsular India (http://flora-peninsula-indica.ces.iisc.ac.in/welcome.php; accessed on 31 May 2020), India Biodiversity Portal and Flowers of India (http://flowersofindia.net/; accessed on 31 May 2020), were checked to extract general information of the species. Further, global databases like the database of the Natural Resources Conservation Service (NRCS) (https://www.nrcs.usda.gov/wps/portal/nrcs/site/national/home/; accessed on 28 May 2020) was used to obtain information. For the small number of species, for which data was absent in all of the above databases, Google search engine (https://www.google.com/; accessed on 31 May 2020) was used, and relevant information was retrieved using the following keyword combinations: species name, growth habit and duration (including all categories of growth habit and duration).

Pathway and time of introduction

Residence time is an important predictor of naturalization; the longer a species is present in a region, the more propagules it can produce, resulting in a higher probability of founding new populations [29]. A positive relationship is, therefore, frequently found between the residence time in an introduced range and the naturalization success of an alien species [30]. It is therefore not surprising that identification and prioritization of the introduction pathway of alien invasive species have been recognized as one of the Aichi Biodiversity Targets (Target 9) (www.cbd.int/decisions/cop/?m=cop-10; accessed on 1 June 2020) and retained as one of the targets (Target 5) to be achieved by 2030 in the zero draft of the post-2020 global biodiversity framework (https://www.cbd.int/doc/c/3064/749a/0f65ac7f9def86707f4eaefa/post2020-prep-02-01-en.pdf; accessed on 13 September 2020).

Information regarding the introduction pathway was retrieved from Google Scholar. We used the advanced search function of the Google Scholar database (https://scholar.google.com/#d=gs_asd; accessed on 20 May 2020) to search for a species name with at least one of the search queries related to introduction purpose ["introduction history" OR "introduced" OR "introduction purpose" OR "introduction pathway"] present anywhere in the article with the exact phrase "India". The screening of articles was continued until the relevant information for a specific introduction purpose was obtained for a species. In addition to these sources, we relied on our knowledge of the Indian literature and searched two other journals, namely Tropical Ecology archives and Indian Forester, which are not indexed in the Google Scholar database. Information for the remaining species (for which there was no data available from the literature) was retrieved from CABI-ISC database (accessed on 12 June 2020). If the introduction purpose of the species could not be established for India, then possible

purposes for which the concerned species had spread outside of its native range were identified by searching authoritative literature records. A structured query search string ["species name" AND ("introduction history" OR "introduction pathway" OR "introduction purpose" OR "introduced")] was used to retrieve this information using ISI Web of Science platform (https://webofknowledge.com/; accessed on 13 June 2020). The reasons for introductions were categorized further following the introduction pathway classification framework proposed by the Convention on Biological Diversity [31].

We recorded the introduction time of the species mentioned in the literature. In addition, the oldest occurrence record of the species in the country was also retrieved from – 1) the Global Biodiversity Information Facility (GBIF) (http://www.gbif.org; accessed on 16 June 2020) and 2) the Global Alien Species First Record Database version 1.2 [5]. The GBIF database was suitable for this exercise due to its comprehensive wealth of information on museum specimens collected in the 18th and 19th century whereas the Global Alien Species First Record Database holds region specific first records of 47542 alien established taxa. To assess the residence status (sensu [21]) of the species in India, we considered the earliest year recorded among literature and these two databases. Finally, the information on the native range for each of these species was retrieved from GRIN database (accessed on 4 April 2020) following the continental and regional scheme (Level 2) proposed by the International Working Group on Taxonomic Databases for Plant Sciences (TDWG) [32].

Global and Indian naturalized ranges

More than 13,000 vascular plant species have already become naturalized outside of their native ranges [4], and the naturalization risk of alien species is predicted to increase due to climate change [33]. Further, a recent study has reported that warm temperate and sub-tropical regions have a higher relative richness of naturalized alien plant species [34]. In this context, information on the geographical range, where each of the species has become naturalized across 1029 global regions, was recorded from the GloNAF database. The naturalized ranges of the species were recorded by identifying the subset of geographical areas that fall within the political boundary of India (region_id: 380-402 in the GloNAF database). This provided us with the information on global as well local (nationwide) naturalized ranges for each of the species. The finest resolution of naturalized range information provided by GloNAF is TDWG-Level 4, corresponding to the constituent political units of a country, i.e., states and union territories of India. The TDWG-Level 4 range information for the other species (i.e., except the naturalized ones) was retrieved from the CABI-ISC database (accessed on 12 June 2020).

Economic uses and market dynamics

Economic activities have the potential of acting as direct [35, 36] as well as indirect drivers [37, 38] of plant invasion. The synchronous increase in trade, transport, and tourism act as a major driving force behind the increasing alien species introductions across the globe [39, 40]. Additionally, most of the intentionally introduced plant species tend to have multiple economic uses, e.g., cultivation, ornamental, medicinal uses, which can increase the propagule pressure and promote naturalization success [41]. To capture such economic usage information for the alien flora of India, we first extracted information from the book Ethnobotany of India [42]. The five volumes of this book provided detailed and updated information of plant uses focusing on major biogeographical zones of the country. The medicinal use in India was recorded in binary format (Present/Absent) by surveying the online database of the Environmental Information System's Centre on Medicinal Plants (http://envis.frlht.org/; accessed on 14 June 2020), operational under the aegis of the Ministry of

Environment, Forests and Climate Change (MoEFCC), Government of India. Two additional databases, namely GRIN and CABI-ISC (accessed on 12 June 2020), were also explored. These databases provided information on economic uses reported across the distribution range of each species. Also, Useful Tropical Plants (http://tropical.theferns.info/; accessed on 15 June 2020) and Useful Temperate Plants (http://temperate.theferns.info/; accessed on 15 June 2020) databases were probed to document the ornamental uses. The thirteen level 1 states of plant uses as prescribed by TDWG in their Economic Botany Data Collection Standard [43] was then used as a template to categorize the uses for each of the species.

Ornamental trading of species is known as one of the major drivers for the introduction and spread of alien plant species worldwide [35, 44, 45]. To comprehend the trading scenarios of these species in India for ornamental purposes and their corresponding translation to revenue streams, two e-commerce platforms were scouted. In India, the e-commerce industry has been on an upward growth trajectory and is recognized as the fastest growing industry in India (https://www.ibef.org/industry/ecommerce.aspx, accessed on 10 July 2020). Further, e-commerce trade, being devoid of strict bio-security regulations, is providing a new distribution channel for alien and invasive species worldwide and India is no exception [46]. For this exercise, we considered two online nurseries, namely Nurserylive (https://nurserylive.com) and Plantslive (https://www.plantslive.in), as the representatives of online ornamental plant trading scenario in India because they – i) mention sales history of each plant species, ii) provide national-level shipping facilities, and iii) include the scientific names of each species in their record. Data collected from these portals included types of plant material being sold (seed/plant), quantities of each species sold during the previous month (i.e., June 2020; retrieved on 15 July 2020), and their unit price (in INR).

Geographic distribution

Information on the current distribution of alien species in a geographical landscape is important to understand the ecological and environmental interactions shaping the introduction to naturalization continuum and has a broad range of applications regarding the management of naturalized and invasive species [47]. To estimate the geographical distribution of the species across the country, their occurrence records (latitude and longitude) species were extracted from GBIF (accessed on 5 July 2020; https://doi.org/10.15468/dl.7bkqza). Since the web interface allows downloading of occurrence data for a maximum of 200 species, we used the *rgbif* [48] and *taxize* [49] packages in R to search and download the occurrence records by matching the taxon keys with GBIF and by using the following predictors: country (India), coordinate (yes), geospatial issue (false). The GBIF database was searched with the standardized taxon names along with their corresponding author names (see ‘Categorization of the alien flora of India' section). The occurrence records were cleaned thereafter to align the data with the mother database. First, the points falling in the ocean were removed from the dataset. Second, as GBIF retrieves occurrences using synonymous species names, the occurrences of the synonymous species were merged. Similarly, occurrence records were merged for the species that had more than one author names, and the author’s name was retained according to the mother database. Finally, the occurrence data of the infraspecific taxa were removed to align the occurrence dataset with the species-level mother database. The occurrence records were then used to determine the presence of each of the species across the different land-use and land-cover (LULC) classes, ecoregions, and anthropogenic biomes.

To identify the species' distribution across the different LULC classes, we used a recently classified image of the country having a spatial resolution of 56 m acquired by the multi-season IRS Resourcesat-2 Advanced Wide Field Sensor (AWiFS) [50]. The classified image identified land-use

and land-cover classes based on natural vegetation types (excluding cultivated systems, croplands, plantations, barren land, water bodies, and snow). By using the spatial extraction tools in ArcMap 10.2.1, LULC information for each of the occurrence records was extracted. This information was then processed in R to identify the presence or absence of the species in 34 LULC classes. Notably, LULC is a dynamic variable and information of LULC classes occupied by a species should be treated as the potential LULC classes which the species can occupy and in no way reflects its current distribution.

The species' occurrences were further determined across different ecoregions and anthropogenic biomes. An ecoregion is characterized as land having a distinct assemblage of species and natural communities with boundaries *a priori* of major land-use change [51]. The anthropogenic biomes, also known as anthromes, characterize anthropogenically-altered terrestrial biosphere by using patterns generated from human-ecosystem interactions [52]. Out of the 814 global ecoregions, 51 were found in India distributed within 12 biomes and the Indo-Malay biogeographic realm. Out of the 21 global anthromes, 18 anthromes were found in India. By using the intersection and spatial extraction tools in ArcMap 10.2.1, we identified the species' distribution across the different ecoregions and anthromes present in the country using a custom script in R.

In addition, the information regarding the habitats of the species was also collected from CABI-ISC (accessed on 12 June 2020). The habitat information was extracted at two levels: the coarser system level (terrestrial, freshwater, littoral and brackish) and the finer sub-system level (e.g., forests, riverbanks, wetlands, deserts, urban areas).

Climatic niche

The bioclimatic variables, derived from the monthly temperature and rainfall values, are considered as biologically meaningful variables and are used in various ecological modelling studies [53]. Occurrence records were used to characterize the realized climatic niche of the species for two bioclimatic variables - annual mean temperature (BIO1) and annual precipitation (BIO12). The bioclimatic variables (averaged over 1970-2000) were downloaded from WorldClim version 2 at a spatial resolution of 2.5 arc minutes (approximately 5 km resolution at the equator) [54]. The BIO1 and BIO12 values were extracted for each of the occurrence records using the Spatial Analyst Toolbox of ArcMap 10.2.1. The extracted information was further processed in R to identify the range (minimum and maximum) and average (±SD) annual temperature and precipitation values for each species. Notably, temperature and precipitation values should be interpreted as realized niche of a species (i.e., from where the species has been recorded yet). The fundamental niche of the species is ‘larger’ than the realized niche [55] and can only be detected when the species colonizes new environmental conditions by overcoming competitive interactions and other limiting factors which are responsible for restricting the realized niche of the species [56].

The distribution of the species across different climatic zones in India was further identified. For this purpose, we used the Köppen-Geiger empirical climate classification system, which classifies the global climate into five major classes and 30 sub-types [57]. The species' occupancy of the climate classes present in India was ascertained using the *kgc* package [58] in R, in which occurrence records were used as input data.

Assessment of data quality

The data quality was assessed by using a checklist having three criteria – reliability, usability, and accessibility. The individual data records of the dataset were extracted from highly-cited and

well-regarded databases, peer-reviewed authoritative scientific journals, and published books. Therefore, the curated data should be considered as reliable and consistent. However, there are many sources of uncertainty in such data [59]. We further adopted a two-step process for technical validation of the curated data. Intense discussions among the developer team were undertaken to maintain the quality of the adopted strategies and techniques of data collection. First, the citations and sources provided by the larger databases were randomly checked to acknowledge the quality of the data provided. Hence, the accuracy of the data sources was successfully established. Second, after completion of data curation for each variable, data for 10% species were rechecked by another curator of the team. The mismatches between data collected by two curators, if any, were resolved and the rechecking process was repeated with 15% species. The process was repeated with an increment of 5% species at every turn until there was no variation between curators' observations. Third, duplicate data extraction was performed for 10% species (including species with missing values) to remove exclusion bias (i.e., to avoid omission errors). Finally, before including the data in the mother database, profiles of 10% species (n = 175) were rechecked by an individual team member. At this stage, we noted no mismatch between individual team member's observations, thereby reliability of the curated data was assured.

To ensure smooth usability of the database by the end users, we arranged individual variables in separate .csv files with keeping species names consistent across the data files. We further included a document in the database as a standalone file explaining the column heads for each data file separately. Finally, the data files were made accessible without any restriction of use through an online data repository and a dedicated website, the details of which have been provided in the 'Data accessibility' section.

Results and Discussion

Curating data from 22 sources, including both national and global databases, ILORA version1.0 incorporates attributes of 1747 alien vascular plant species of India in thirteen data variables. Of these, two variables, namely the species categorization and general information, contain data for all the species (Fig 2A). The database has recorded common names for 1575 species and vernacular names for 583 species in 23 languages. For the remaining 11 data variables, 1671 species have information for one or more variables whereas data is missing for 76 species (Fig 2B). Among the three categories of alien flora (invasive, naturalized, and casual), average data availability is maximum for invasive aliens followed by naturalized and casual aliens (Fig 2C).

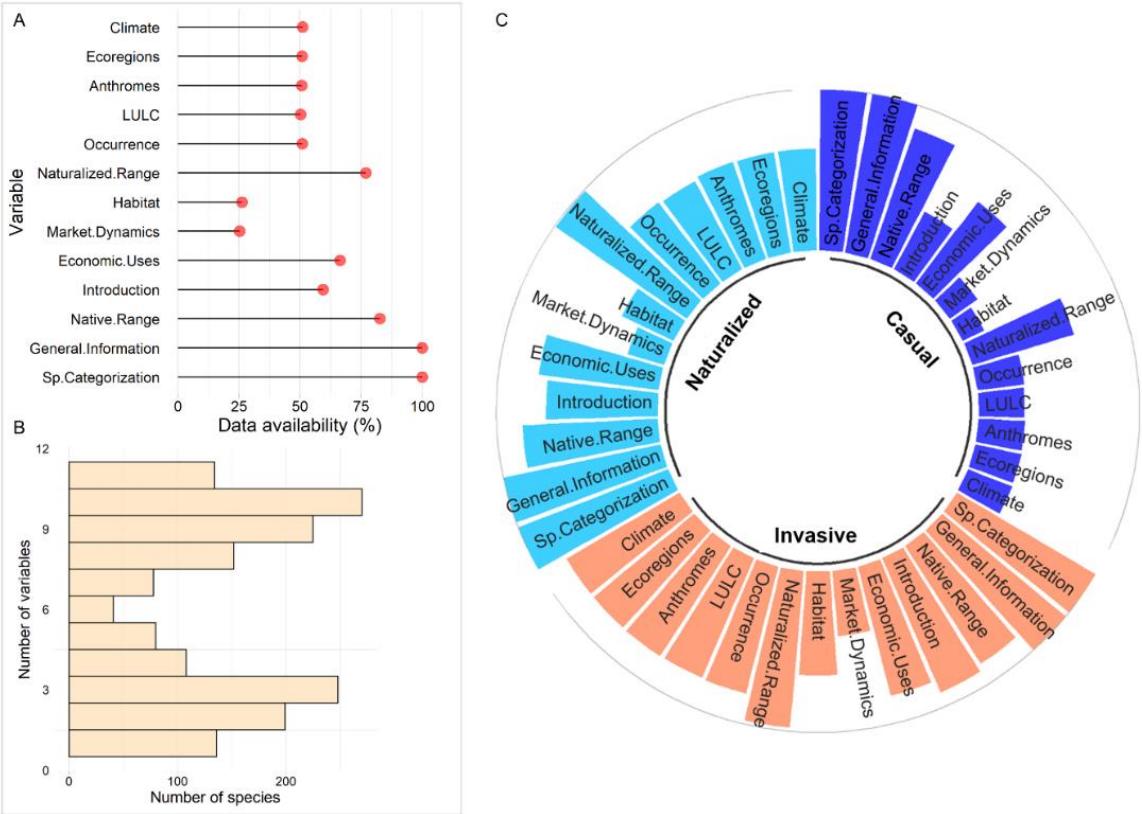

Fig 2. Data availability in the ILORA database – (A) at the variable level: percentage of available data for each of the 13 variables included in the database; (B) at the species level: the frequency histogram showing numbers of species (n=1671 having data for at least one variable) for each of the 11 variables (i.e., excluding two variables, species categorization and general information having no missing data); (C) at the category level: the circular bar plot showing the percentage (the grey circle outside represent 100% data availability) of available data for each of the 13 variables for three categories of alien species (invasive, naturalized and casual). The plots have been prepared using the *ggplot2* package [60] in R.

At the taxonomic level, the alien species (invasive, naturalized, and casual) data are distributed across 764 genera and 154 families with Asteraceae, Fabaceae and Poaceae as the most represented families (Fig 3A). Out of the 764 genera, 529 genera have species belonging to either one of the three categories (invasive = 76, naturalized = 72, casual = 481) whereas 21 genera have species of all three categories (Fig 3B, 3C).

The data records for the species identify 50 native ranges (at TDWG Level 2 which correspond to subcontinental regions) for 1444 species and 919 naturalized ranges (which correspond to the states, provinces, or administrative regions within countries) for 1344 species. This has been supplemented with information on the possible introduction pathways of the 1038 species to regions beyond their native ranges. Considering the native range information of the three alien species categories (invasive, naturalized, casual), the Americas have been found to be the native range of most of these species followed by Africa and Asia-Temperate whereas the introduction for ornamental purpose other than horticulture has been found to be the dominant introduction pathway (Fig 4).

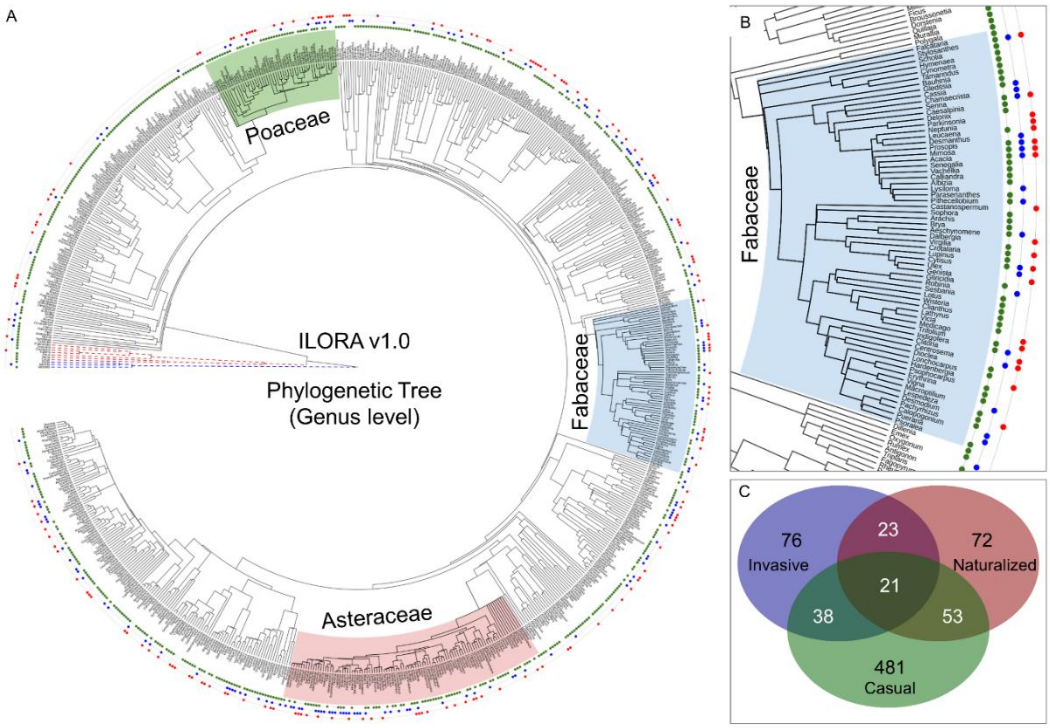

Fig 3. Visualization of taxonomic data of invasive, naturalized, and casual aliens – (A) Genus-level phylogenetic tree showing phylogeny of the 764 genera belonging to 154 families. The dotted branches represent gymnosperm (red) and aquatic angiosperm (blue) genera, the three families having maximum number of taxa have been highlighted. The phylogenetic tree was visualized and annotated using the web-based tool Interactive Tree Of Life v.3 [61]; (B) The green dots indicate genera having casual alien species, blue dots for invasive aliens and red dots for naturalized aliens; (C) The Venn diagram (generated using *VennDiagram* package [62] in R) showing number of genera having casual, naturalized and alien species, either individually or in combination.

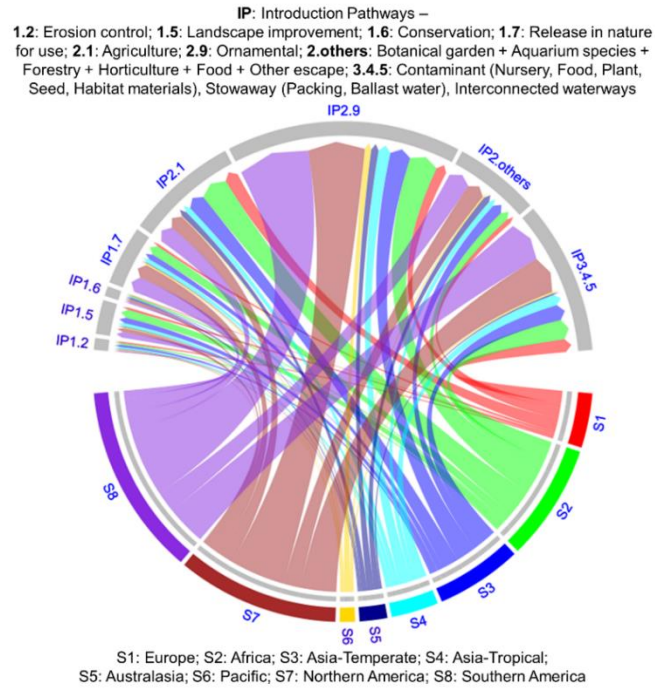

Fig 4. Flows of alien species between native range and introduction pathways. The chord diagram (generated using the *circlize* package [63] in R) shows the number of alien species (n=317, considering invasive, naturalized and casual alien species for which both native range and introduction pathway information are available in the database) introduced from each continent by the different introduction pathways. The width of a chord represents the number of alien species.

Herbaceous growth form has been found to be predominant across all five categories of species (Fig 5A). The data on potential habitats of 459 species identifies 38 subcategories that are combined into five coarser levels. The database is also populated with 40,724 occurrence records for 889 species present across 34 Land-Use and Land-Cover (LULC) classes, 18 anthromes, 42 ecoregions and 15 climate classes in India. The invasive aliens have been found to occupy maximum number of LULC classes, anthromes and ecoregions (Fig 5B). Further, 50 subcategories of uses have been also recorded for 1159 species, which were combined into thirteen Level 1 uses (Fig 5C), among which maximum number of species have environmental uses (Fig 5D). The online nursery data for 442 species reveals higher market price of native and cryptogenic aliens compared to that of invasive, naturalized, and casual aliens (Fig 5E).

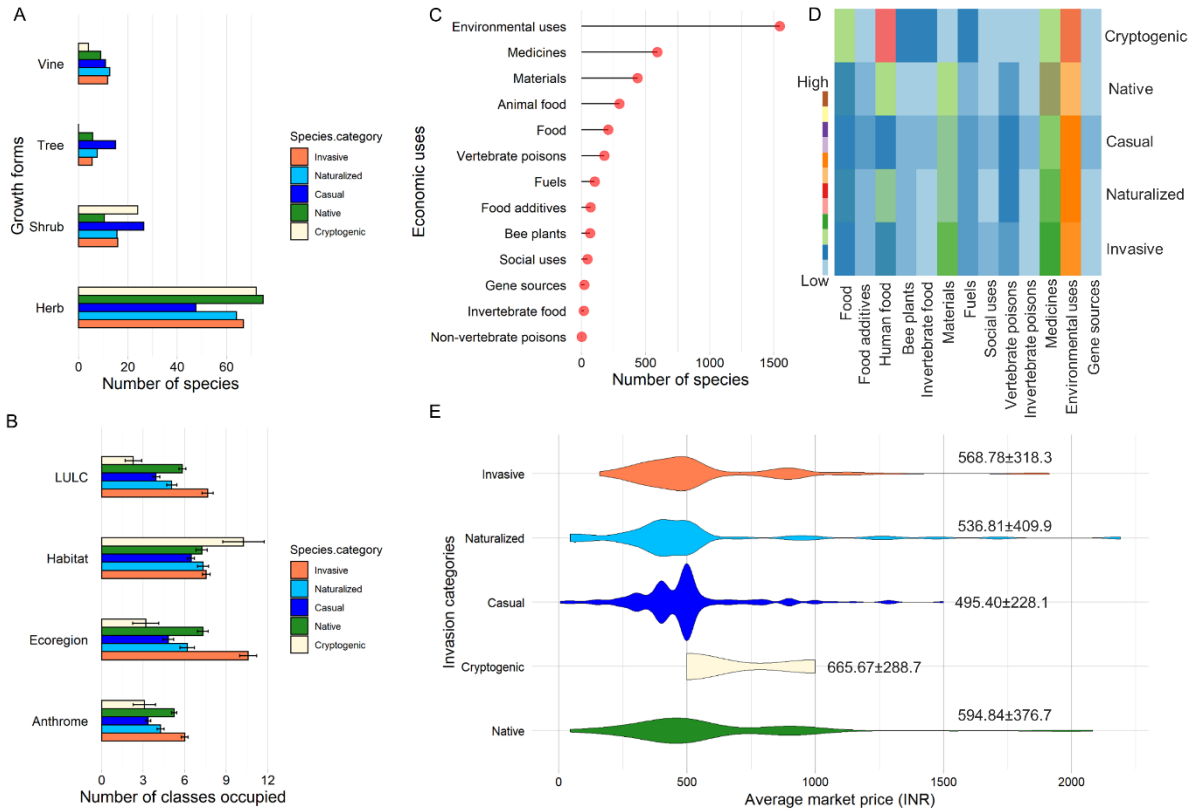

Fig 5. Visualization of variables across five categories of species included in the database – (A) growth form, (B) average (±SE) number of occupied anthromes, ecoregions, habitats and LULC classes, (C) number of economic uses for five categories together, (D) heatmap showing number of economic uses for each of the five categories of species, (E) violin plot showing the distribution of market price values (with mean ± SE values showing above each plot). The plots have been prepared using the *ggplot2* package [60] in R.

Data accessibility

The data records can be accessed through two pathways –

1) Online data repository: The data records are arranged and archived into the 14 files (in .csv format) in the online repository figshare (DOI: 10.6084/m9.figshare.13677460). In each of these files, the records are associated with 'Acc_Species_Name' which allows correspondence across the entire database. The file ILORA_summary.csv contains an overview of the data records available for individual species. Cells with missing or unavailable information are left blank. Four additional files were also included in the data repository – 1) metadata used to determine origin and invasion status

of alien species (in .csv format), 2) a list of cultivated alien species (in .csv format), 3) the R codes used to extract the variable information (in .txt format), and 4) a document explaining the column heads of each of the 14 data files (in .pdf format). An overview of the file name, data contents and data type have been provided in Table 1.

2) Dedicated website: ILORA website (https://ilora2020.wixsite.com/ilora2020) is a user-friendly, public platform that can be used to retrieve necessary data records through query-based search of the database. A RShiny [64] application has been created and embedded on the website for providing easy and complete access of the database in an interactive manner. The data files or a subset of the data (based on query) can be downloaded in .csv format. The scripts used to process the data files are also available in a Github repository (https://github.com/Ilora2021/ILORA/). The website is also supplemented with data highlights, visual representation of data, and associated infographics, which enables the information to permeate a broader range of audiences. It also provides a feedback loop for the users to input new data or update the existing data, thereby contributing to the dynamism and long-term sustainability of the database. As is the common practice for any database, updated information will be communicated to the subscribed users on a regular basis.

Table 1. Overview of the data records of ILORA version 1.0 representing the contents and data type for each of the eighteen data files.

	File name	**Data contents**	**Data type**
1	ILORA_1_SpCategorization.csv	The reported origin and invasion status of the alien species present in India	Nominal
2	ILORA_1a_Sp.Categorization.Metadata.csv	Metadata for determining origin and invasion status of the alien species	Nominal; Binary
3	ILORA_2_GeneralInformation.csv	Life form, duration, and type of the species	Nominal
4	ILORA_3_NativeRange.csv	Native range number (TDWG Level 2) of the species	Nominal
5	ILORA_4_Introduction.csv	Month and year of first reported occurrence in India; possible introduction pathway(s)	Nominal
6	ILORA_5_EconomicUses.csv	Economic uses associated with the species, arranged following TDWG Level 1 scheme	Binary
7	ILORA_6_MarketDynamics.csv	Price and number of plants/seeds sold in two online nurseries of India	Nominal; Binary
8	ILORA_7_Habitat.csv	Potential habitat(s) of the species	Binary
9	ILORA_8_NaturalizedRange.csv	Naturalized ranges of the species across the globe and in India (TDWG Level 4)	Nominal
10	ILORA_9_Occurrence.csv	Latitude and longitude data of the alien plant species	Nominal

11	ILORA_10_LULC.csv	Land-Use and Land-Cover categories of India that are inhabited by the species	Binary
12	ILORA_11_Anthromes.csv	Indian anthromes that are inhabited by the species	Binary
13	ILORA_12_Ecoregions.csv	Indian ecoregions that are inhabited by the species	Binary
14	ILORA_13_ClimateClass.csv	Köppen -Geiger climate classes of India inhabited by the species; and the temperature and precipitation data of the species	Nominal; Binary
15	ILORA_14_Summary.csv	Status (presence/absence) of data for each of the above mentioned 13 data files for each of the alien plant species	Binary
16	ILORA_15_Cultivated.Species.csv	Standardized species names of cultivated aliens	Nominal
17	ILORA_16_R.Codes.txt	R codes used to extract variable information	Document
18	ILORA_HowToRead.pdf	Document having information of the data types recorded in columns in each of the data file	Document

Data usage

The primary objective behind curating ILORA is to provide a comprehensive, easy-to-access, policy, and research-relevant database for alien vascular plant species of India. It contains data of two (alien species occurrence and species alien status) of the three Essential Variables for Invasion Monitoring (except the alien species impact) [65] of alien plant species present in India. ILORA should not be considered as a checklist of alien flora present in India. Rather, the essential and the supplementary variables (i.e., the derived variables like the pathways of introduction and characteristics of the receiving environment), which have been collated both *in-situ* and from *ex-situ* sources [65], provide a baseline data that can be utilized for identifying, reporting, monitoring, and regulating the status of alien plant species in India. With data records ranging from taxonomy, perceived alien status, introduction history to socio-economic uses, and geographical distribution, ILORA envisages assisting the various stakeholders who are associated with alien plant invasion research and policymaking in India.

The data records of ILORA could be incredibly useful to cater to the need of data availability for research on alien plant species in India. For example, the database can be used to identify the relative importance of variables along with the progression of alien species through the introduction-naturalization-invasiveness continuum. The database can be of immense value to identify potential invasive species. A recent study has modelled the impacts of the environment, species characteristics and nursery availability to identify the predictors for naturalization and to forecast naturalization risks of ornamental plant species in Europe under future climate conditions [66]. Similar modelling exercises could be crucial in understanding the dynamics of the alien species and in streamlining the management efforts towards timely containment or eradication, especially for a developing country like India. Further, the broad range of distribution information can be used for habitat suitability,

correlative and mechanistic distribution modelling approaches for predicting the potential distribution and range expansion of naturalized and invasive species as well as of the emerging invaders in the country. These findings may inform policy directives for timely management of the alien plant species based on their residence status across the invasion growth curve. On the other hand, data on introduction purpose and naturalized range of the species could be used for formulating international trading policies, imposing stricter border control measures to prevent novel introductions, as well as developing domestic trading policy interventions to check the further spread and naturalization of species. ILORA also allows users to comprehend the socio-economic drivers of alien plant spread in the country. Similarly, information on the distribution of species across different climate classes, anthromes, ecoregions, and LULC classes could guide the implementation of site-specific management protocols to prevent naturalization of alien species.

ILORA has also provisions for future users to contribute data and information to ensure bidirectional information exchange. Users can submit data for all 13 variables included in version1 of the database for the species already present in the database and also for new species to be included. The standard for such data submission has been given on the database website. For each variable, users need to provide detail information of the data sources which can be traceable and authenticated. Such sources include, but are not limited to, published journal article (preferably with DOI), book or book sections (preferably with DOI), historical text, thesis report, conference paper, published dataset, online database, and government document. In case the submitted data is not verifiable by an authenticated source (e.g., personal observation), users need to submit detailed descriptions of the data collection (e.g., photographs with date and time stamps with the GPS feature enabled). The submitted data will undergo verification and thorough quality-checks by the curators (and domain experts as and when required) before integration in the database to ensure data quality. All data will be associated with source information and ILORA will encourage users to refer back to the primary sources so that the content providers can get full credit and acknowledgement. This practice will allow expansion of the resolution as well as the scope of the database.

Future directions

Sustainability measures

ILORA is a standardized repository containing biological and socio-economic data of alien plant species present in India which can be helpful for a host of research applications, policy directives and decision-making approaches. We envisage that the database can be transformed into a nationwide collaborative platform for plant scientists, bridge the knowledge gaps and assist novel research on alien plant invasions in India. Therefore, we have certain plans for long-term maintenance of ILORA based on funding, collaboration, and voluntary participation. First, we are applying for the necessary funding in India to develop and sustain a dedicated workforce to ensure database update at a regular interval, monitoring the inflow of new information and responding to queries and suggestions of end users. Second, for a wider dissemination and utility of the database as well as to attract user inputs, collaboration opportunities will be explored. Initially, we plan to share the database with the Indian academia who are actively engaged in invasive species research and to integrate the database with the existing repositories (e.g., India Biodiversity Portal). Such collaborations will be further used to contribute data to global databases. For example, India is a GBIF observer country from Asia and the occurrence records of Indian alien flora will be hosted in GBIF in a Darwin Core Archive (DwC-A) format through the coordinating and thematic nodes of the existing national network. Databases focused on invasive alien species at various scales are seldom linked with each other. ILORA has been made in such a way that it can be easily integrated with other global databases

and repositories (e.g., following the Biodiversity Information Standards, aligning the naturalized range information with the GloNAF database and habitat information with CABI-ISC). Regular updates of ILORA based on the availability of new information in global and Indian resources are therefore expected to address the transferability and inter-operability issues across the databases while being operational at national and regional scales. Third, the dynamic version of the database will be equipped with query-based search and information retrieval functions as well as a simplified and structured data submission system to enhance end user experience and attract data submission. These automated features along with the provision of citing the original sources and thereby acknowledging the contributor are expected to attract voluntary participation and ensure the long-term sustainability of the database.

Qualitative and quantitative enhancements

Alien species databases are not complete and stable units; dynamism lies at the core of the data records and thus, they are open to suggestions and future updates. ILORA is also not exhaustive in every aspect as there is a lot of scope for the enrichment of data resolution. The subsequent versions of the database will also increase data resolution by including – i) detail profiles of the cultivated aliens (since many of these species may escape human control and become naturalized; all data files), ii) potential dispersal pathways (valuable for the prevention of further spread and reduce chances of novel invasion; ILORA_Introduction.csv), iii) trading scenarios (both online and offline sales figures of alien plant species in the market which can be used as a proxy for propagule pressure; ILORA_MarketDynamics.csv), and iv) functional traits (following the leaf-height-seed scheme proposed by [67]; ILORA_GeneralInformation.csv).

While we, as the developers, will work towards increasing the quantitative and qualitative enhancements of ILORA, we have also identified the following scopes which grant opportunities for contributions from the scientific community, citizens, and stakeholders for further improvement of the data quality. First, the origin status of some of the alien species was found to be contradictory between the national checklists and global databases. Such discrepancies may arise from knowledge gaps between respective curators leading to lack-of-consensus decisions; nevertheless, it can be resolved by the thorough and scientifically-informed pursuit of the literature and archival records. However, this exercise is often challenging, given that many such records are not systematically digitized or included in searchable databases. One possible way around this is to rely on in-depth biogeographic knowledge accumulated from the working experience of the concerned taxa. Further, given the large geographic extent of the country, it is likely that some species are native to one part of India but naturalized in other parts. For example, *Nasturtium officinale* has been recognized as naturalized in India across the three databases [11, 13, 17]. According to POWO, the species is native in Pakistan whereas GRIN recognizes the species as native in India. If pre-partitioned British India is considered, the species is likely to be native in northern India (a politically different entity from Pakistan, but biogeographically similar) from where it might have been introduced to peninsular India and became naturalized (personal communication). Therefore, the resolution of origin and invasion status of the alien species can be increased many folds by considering documented introduction history and robust biogeographical evidence, if available, on a case-by-case basis (ILORA_Sp.Categorization.csv). Similarly, the majority of the invasive plant species reported in different countries in South Asia were reported or imported first in India. Initiatives similar to ILORA by other South Asian countries, or preferentially a common database of this kind, will be an important step towards preventing and managing invasive alien species at the regional level. By providing the baseline information available for each species, ILORA opens the door for further discussion among the scientific community and

expects knowledge support from the researchers, especially from the taxonomists and phylogeographers, to refine the database.

Second, while the database curated the available information on the introduction pathway at the highest available resolution, we are aware that detailed scientific studies of a particular taxon can generate novel information or modify the existing knowledge. For example, *Mikania micrantha* (Compositae) has been long known to be introduced during the 2nd World War to camouflage the airfields of northeast India [68]. However, reconstruction of the introduction pathways based on herbarium records and archival literature revealed multiple introductions of the species in different regions in India [69]. Similar fine-scale information is only available for a few species, primarily due to focused research interest of Indian academia and unavailability of archival and herbarium records in a digital and easy-to-access format. ILORA, therefore, welcomes the contribution of scientific knowledge for species with unknown introduction pathways and modification of the existing information, if any, by presenting evidence (ILORA_Introduction.csv).

Third, the resolution of the occurrence records as well as of the distribution information can be improved by many folds. Presently, the occurrence records were extracted from GBIF due to its comprehensive coverage of both historical herbarium collections as well as contemporary citizen contributions. However, the occurrence records in GBIF are not exhaustive and often poorly represent the actual distribution of a species. For example, for a widespread species like *M. micrantha*, GBIF has only 15 records from India which can be certainly improved by including personal observations, research data and herbaria information. If collected, this information would help in achieving a higher data resolution, as demonstrated by the FLOTROP database on plant diversity of northern Africa by including locally available information with the existing data in GBIF [70]. While we will work towards curating occurrence records from the published literature (ILORA_Occurrence.csv), it may not be possible for us to capture the vast information stored in local herbaria and reports in vernacular languages. The occurrence data stored in our database would surely be valuable for broad-scaled research and applications, and contributions from future users of ILORA of hitherto unavailable information could significantly increase the resolution of the database.

Acknowledgements

We express our sincere gratitude to our respective organizations, institutions, and universities for giving us the logistic support, which was an essential requirement in the absence of any specific grant for conducting this research. We acknowledge the support of Dr. C. Sudhakar Reddy from National Remote Sensing Centre (Indian Space Research Organization, Hyderabad) who provided us with the classified image of the country. We are thankful to the editor and four reviewers (Prof. Dr. Mark van Kleunen of the Konstanz University, Germany; Prof. John Ross Wilson of the Stellenbosch University, South Africa; Dr. Bharat Babu Shrestha of Tribhuvan University, Nepal; and one anonymous reviewer) for their valuable comments and constructive criticisms which immensely improved the manuscript and database quality and provided us with enough encouragement and future prospects to continue our work and sustain the database. Global Covid-19 pandemic situation tested our capacities to work on this database and manage families simultaneously; therefore, we would like to specially thank our better halves and children (Sushmita Bhattacharya and Addiriti Mukherjee; Kavita Parab and Tanay Bhowmick; and Payal Banerjee) for their patience and constant support for us.

Author Contributions

Conceptualization: Abhishek Mukherjee, Achyut Kumar Banerjee.
Data Curation: Vidushi Pant, Chinmay Patwardhan, Achyut Kumar Banerjee
Investigation: Vidushi Pant, Abhishek Mukherjee, Amiya Ranjan Bhowmick, Achyut Kumar Banerjee
Project Administration: Achyut Kumar Banerjee
Resources: Abhishek Mukherjee, Amiya Ranjan Bhowmick, Achyut Kumar Banerjee
Software: Kshitij Patil, Amiya Ranjan Bhowmick
Supervision: Achyut Kumar Banerjee
Validation: Vidushi Pant, Abhishek Mukherjee, Amiya Ranjan Bhowmick, Achyut Kumar Banerjee
Visualization: Vidushi Pant, Kshitij Patil, Achyut Kumar Banerjee
Writing – Original Draft Preparation: Vidushi Pant
Writing – Review & Editing: Amiya Ranjan Bhowmick, Abhishek Mukherjee, Chinmay Patwardhan, Achyut Kumar Banerjee

References

1. Ministry of Environment and Forests. India's fifth national report to the Convention on Biological Diversity. New Delhi, India: Ministry of Environment and Forests, Government of India, New Delhi, India; 2014.

2. Scarrow R. Botanical Imperialism. Nature Plants. 2018;4(6):316-. doi: 10.1038/s41477-018-0173-2.

3. Bhagwat SA, Breman E, Thekaekara T, Thornton TF, Willis KJ. A battle lost? Report on two centuries of invasion and management of *Lantana camara* L. in Australia, India and South Africa. PLOS ONE. 2012;7(3):e32407. doi: 10.1371/journal.pone.0032407.

4. van Kleunen M, Dawson W, Essl F, Pergl J, Winter M, Weber E, et al. Global exchange and accumulation of non-native plants. Nature. 2015.

5. Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM, et al. No saturation in the accumulation of alien species worldwide. Nature communications. 2017;8:14435.

6. Lenzen M, Moran D, Kanemoto K, Foran B, Lobefaro L, Geschke A. International trade drives biodiversity threats in developing nations. Nature. 2012;486(7401):109-12. doi: 10.1038/nature11145.

7. Seebens H, Essl F, Dawson W, Fuentes N, Moser D, Pergl J, et al. Global trade will accelerate plant invasions in emerging economies under climate change. Global Change Biol. 2015;21(11):4128-40. doi: doi:10.1111/gcb.13021.

8. Pagad S, Genovesi P, Carnevali L, Schigel D, McGeoch MA. Introducing the Global Register of Introduced and Invasive Species. Scientific Data. 2018;5(1):170202. doi: 10.1038/sdata.2017.202.

9. Chatterjee D. Studies on the endemic flora of India and Burma. Journal of Asiatic Society of Bengal. 1940;5:19-67.

10. Nayar MP. Changing patterns of the Indian flora. Bull Bot Surv India. 1977;19:145-55.

11. Khuroo AA, Reshi ZA, Malik AH, Weber E, Rashid I, Dar G. Alien flora of India: taxonomic composition, invasion status and biogeographic affiliations. Biol Invasions. 2012;14(1):99-113.

12. Sankaran KV, Khuroo A, Raghavan R, Molur S, Kumar B, Wong LJ, et al. Global Register of Introduced and Invasive Species - India. Version 1.3. Accessed via GBIForg on 2020-05-13; https://doiorg/1015468/uvnf8m. 2020.

13. van Kleunen M, Pyšek P, Dawson W, Essl F, Kreft H, Pergl J, et al. The Global Naturalized Alien Flora (GloNAF) database. Ecology. 2019;100(1):e02542. doi: 10.1002/ecy.2542.

14. Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM, et al. Global rise in emerging alien species results from increased accessibility of new source pools. Proc Natl Acad Sci U S A. 2018;115(10):E2264-E73. doi: 10.1073/pnas.1719429115.

15. Faulkner KT, Spear D, Robertson MP, Rouget M, Wilson JRU. An assessment of the information content of South African alien species databases. Bothalia - African Biodiversity & Conservation. 2015;45:1-11.

16. Pyšek P, Lambdon PW, Arianoutsou M, Kühn I, Pino J, Winter M. Alien Vascular Plants of Europe. Handbook of Alien Species in Europe. Dordrecht: Springer Netherlands; 2009. p. 43-61.

17. Inderjit, Pergl J, van Kleunen M, Hejda M, Babu CR, Majumdar S, et al. Naturalized alien flora of the Indian states: biogeographic patterns, taxonomic structure and drivers of species richness. Biol Invasions. 2018;20(6):1625-38. doi: 10.1007/s10530-017-1622-y.

18. Kindt R. WorldFlora: An R package for exact and fuzzy matching of plant names against the World Flora Online taxonomic backbone data. Applications in Plant Sciences. 2020;8(9):e11388. doi: https://doi.org/10.1002/aps3.11388.

19. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.

20. Borsch T, Berendsohn W, Dalcin E, Delmas M, Demissew S, Elliott A, et al. World Flora Online: Placing taxonomists at the heart of a definitive and comprehensive global resource on the world's plants. Taxon. 2020;69(6):1311-41. doi: https://doi.org/10.1002/tax.12373.

21. Pyšek P, Richardson DM, Rejmánek M, Webster GL, Williamson M, Kirschner J. Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. Taxon. 2004;53(1):131-43. doi: 10.2307/4135498.

22. Groom Q, Desmet P, Reyserhove L, Adriaens T, Oldoni D, Vanderhoeven S, et al. Improving Darwin Core for research and management of alien species. Biodiversity Information Science and Standards. 2019;3:e38084. doi: 10.3897/biss.3.38084.

23. Turbelin AJ, Malamud BD, Francis RA. Mapping the global state of invasive alien species: patterns of invasion and policy responses. Global Ecol Biogeogr. 2017;26(1):78-92. doi: 10.1111/geb.12517.

24. Martin PA, Newton AC, Bullock JM. Impacts of invasive plants on carbon pools depend on both species' traits and local climate. Ecology. 2017;98(4):1026-35. doi: 10.1002/ecy.1711.

25. Essl F, Bacher S, Genovesi P, Hulme PE, Jeschke JM, Katsanevakis S, et al. Which taxa are alien? Criteria, applications, and uncertainties. Bioscience. 2018;68(7):496-509. doi: 10.1093/biosci/biy057.

26. Blackburn TM, Pyšek P, Bacher S, Carlton JT, Duncan RP, Jarošík V, et al. A proposed unified framework for biological invasions. Trends Ecol Evol. 2011;26(7):333-9. doi: 10.1016/j.tree.2011.03.023.

27. Jin Y, Qian H. V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. Ecography. 2019;42(8):1353-9. doi: 10.1111/ecog.04434.

28. Kattge J, Bönisch G, Díaz S, Lavorel S, Prentice IC, Leadley P, et al. TRY plant trait database–enhanced coverage and open access. Global Change Biol. 2020;26(1):119-88.

29. Křivánek M, Pyšek P, Jarošík V. Planting history and propagule pressure as predictors of invasion by woody species in a temperate region. Conserv Biol. 2006;20(5):1487-98.

30. Feng Y, Maurel N, Wang Z, Ning L, Yu FH, van Kleunen M. Introduction history, climatic suitability, native range size, species traits and their interactions explain establishment of Chinese woody species in Europe. Global Ecol Biogeogr. 2016;25(11):1356-66.

31. Harrower C, Scalera R, Pagad S, Schonrogge K, Roy H. Guidance for interpretation of CBD categories on introduction pathways. Technical note prepared by IUCN for the European Commission. 2018.

32. Brummitt RK, Pando F, Hollis S, Brummitt N. World geographical scheme for recording plant distributions. Pittsburgh, PA: Hunt Institute for Botanical Documentation, Carnegie Mellon University; 2001.

33. Dullinger I, Wessely J, Bossdorf O, Dawson W, Essl F, Gattringer A, et al. Climate change will increase the naturalization risk from garden plants in Europe. Global Ecol Biogeogr. 2017;26(1):43-53. doi: 10.1111/geb.12512. PubMed PMID: 28111525; PubMed Central PMCID: PMC5216452.

34. Essl F, Dawson W, Kreft H, Pergl J, Pyšek P, Van Kleunen M, et al. Drivers of the relative richness of naturalized and invasive plant species on Earth. AoB PLANTS. 2019;11(5):plz051-plz. doi: 10.1093/aobpla/plz051. PubMed PMID: 31636882.

35. Dehnen-Schmutz K, Touza J, Perrings C, Williamson M. A century of the ornamental plant trade and its impact on invasion success. Divers Distrib. 2007;13(5):527-34.

36. Faulkner KT, Robertson MP, Rouget M, Wilson JRU. Understanding and managing the introduction pathways of alien taxa: South Africa as a case study. Biol Invasions. 2016;18(1):73-87. doi: 10.1007/s10530-015-0990-4.

37. Liu X, Blackburn TM, Song T, Li X, Huang C, Li Y. Risks of biological invasion on the Belt and Road. Curr Biol. 2019;29:499-505.

38. Ward SF, Taylor BS, Dixon Hamil K-A, Riitters KH, Fei S. Effects of terrestrial transport corridors and associated landscape context on invasion by forest plants. Biol Invasions. 2020;22(10):3051-66. doi: 10.1007/s10530-020-02308-3.

39. Hulme PE. Trade, transport and trouble: managing invasive species pathways in an era of globalization. J Appl Ecol. 2009;46(1):10-8.

40. Anderson LG, Rocliffe S, Haddaway NR, Dunn AM. The role of tourism and recreation in the spread of non-native species: A systematic review and meta-analysis. PloS one. 2015;10(10):e0140833-e. doi: 10.1371/journal.pone.0140833. PubMed PMID: 26485300.

41. van Kleunen M, Xu X, Yang Q, Maurel N, Zhang Z, Dawson W, et al. Economic use of plants is key to their naturalization success. Nature Communications. 2020;11(1):3201. doi: 10.1038/s41467-020-16982-3.

42. Pullaiah T, Krishnamurthy K, Bahadur B. Ethnobotany of India. India: CRC Press; 2016.

43. Cook FE. Economic botany data collection standard - Prepared for the International Working Group on Taxonomic Databases for Plant Sciences (TDWG). Richmond, UK: Royal Botanic Gardens (Kew); 1995.

44. Niemiera AX, Von Holle B. Invasive plant species and the ornamental horticulture industry. Management of Invasive Weeds: Springer; 2009. p. 167-87.

45. van Kleunen M, Essl F, Pergl J, Brundu G, Carboni M, Dullinger S, et al. The changing role of ornamental horticulture in alien plant invasions. Biol Rev. 2018;93(3):1421-37. doi: 10.1111/brv.12402.

46. Humair F, Humair L, Kuhn F, Kueffer C. E-commerce trade in invasive plants. Conserv Biol. 2015;29(6):1658-65.

47. Zhou Q, Wang Y, Li X, Liu Z, Wu J, Musa A, et al. Geographical distribution and determining factors of different invasive ranks of alien species across China. Sci Total Environ. 2020;722:137929. doi: https://doi.org/10.1016/j.scitotenv.2020.137929.

48. Chamberlain SA, Ram K, Barve V, Mcglinn D. rgbif: Interface to the Global Biodiversity Information Facility APIR package version 0.8.8 http://CRAN.R-project.org/package=rgbif 2016.

49. Chamberlain SA, Szöcs E. taxize: taxonomic search and retrieval in R. F1000Research. 2013;2:191.

50. Reddy CS, Jha CS, Diwakar PG, Dadhwal VK. Nationwide classification of forest types of India using remote sensing and GIS. Environ Monit Assess. 2015;187(12):777. doi: 10.1007/s10661-015-4990-8.

51. Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GV, Underwood EC, et al. Terrestrial Ecoregions of the World: A New Map of Life on Earth A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. Bioscience. 2001;51(11):933-8.

52. Ellis EC, Ramankutty N. Putting people in the map: anthropogenic biomes of the world. Front Ecol Environ. 2008;6(8):439-47.

53. Jeschke JM, Strayer DL. Usefulness of bioclimatic models for studying climate change and invasive species. Ann N Y Acad Sci. 2008;1134(1):1-24. doi: 10.1196/annals.1439.002.

54. Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology. 2017;37(12):4302-15. doi: 10.1002/joc.5086.

55. Hutchinson GE. Concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology. 1957;22:415-27.

56. Soberón J, Arroyo-Peña B. Are fundamental niches larger than the realized? Testing a 50-year-old prediction by Hutchinson. PLOS ONE. 2017;12(4):e0175138. doi: 10.1371/journal.pone.0175138.

57. Köppen W. Das geographische System der Klimate. Berlin, Germany1936. 1-44 p.

58. Bryant C, Wheeler NR, & RF, French RH. kgc: Koeppen-Geiger Climatic Zones R package version 1.0.0.2. 2017.

59. McGeoch MA, Spear D, Kleynhans EJ, Marais E. Uncertainty in invasive alien species listing. Ecol Appl. 2012;22(3):959-71. doi: doi:10.1890/11-1252.1.

60. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, USA: Springer-Verlag; 2016.

61. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(W1):W242-W5. doi: 10.1093/nar/gkw290.

62. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics. 2011;12(1):35. doi: 10.1186/1471-2105-12-35.

63. Gu Z, Gu L, Eils R, Schlesner M, Brors B. *circlize* implements and enhances circular visualization in R. Bioinformatics. 2014;30(19):2811-2. doi: 10.1093/bioinformatics/btu393.

64. Chang W, Cheng J, Allaire J, Xie Y, McPherson J. shiny: Web Application Framework for R. R package version 1.4.0.2. https://cranr-projectorg/package=shiny. 2020.

65. Latombe G, Pyšek P, Jeschke JM, Blackburn TM, Bacher S, Capinha C, et al. A vision for global monitoring of biological invasions. Biol Conserv. 2017;213:295-308. doi: 10.1016/j.biocon.2016.06.013.

66. Haeuser E, Dawson W, Thuiller W, Dullinger S, Block S, Bossdorf O, et al. European ornamental garden flora as an invasion debt under climate change. J Appl Ecol. 2018;55(5):2386-95. doi: doi:10.1111/1365-2664.13197.

67. Westoby M. A leaf-height-seed (LHS) plant ecology strategy scheme. Plant Soil. 1998;199(2):213-27. doi: 10.1023/A:1004327224729.

68. Tripathi R, Khan M, Yadav A. Biology of *Mikania micrantha* HBK: a review. Invasive alien plants: An ecological appraisal for the Indian subcontinent. Wallingford, UK: CAB International; 2012. p. 99-107.

69. Banerjee AK, Mukherjee A, Guo W, Liu Y, Huang Y. Spatio-temporal patterns of climatic niche dynamics of an invasive plant *Mikania micrantha* Kunth and its potential distribution under projected climate change. Frontiers in Ecology and Evolution. 2019;7:291. doi: 10.3389/fevo.2019.00291.

70. Taugourdeau S, Daget P, Chatelain C, Mathieu D, Juanes X, Huguenin J, et al. FLOTROP, a massive contribution to plant diversity data for open ecosystems in northern tropical Africa. Scientific Data. 2019;6(1):118. doi: 10.1038/s41597-019-0120-8.