INTRODUCTION

1.1 Crop wild relatives

Crop wild relatives (CWR) are the close genetic relatives of domesticated crops, including their progenitors. In addition to providing unique ecosystem functions and biotic interactions in their native environments, CWR represent key sources of genetic material for introduction into crop lines through plant breeding. The use of CWR by agricultural scientists has become regular practice since the 1940s (Meilleur & Hodgkin, 2004), and has contributed to the development of new lines of many globally important crops (Dempewolf et al., 2017; Hajjar & Hodgkin, 2007). Recently, CWR have been included in the tools used to increase the range of conditions in which...
crops can be grown, as well as bolstering adaptability to changing climatic conditions and pathogens (Dempewolf et al., 2017). The importance of these plants for future crop development is highlighted by their specific inclusion as conservation priorities in the Convention on Biological Diversity’s Aichi Biodiversity Target 13 and the United Nations’ Sustainable Development Goals (Target 2.5; United Nations, 2019); and the Convention on Biological Diversity’s Aichi Biodiversity Targets (Target 13: CBD, 2019), Global Strategy for Plant Conservation (GSPC; CBD, 2010) and upcoming Post-2020 Biodiversity Framework (CBD, 2020).

Despite their current and potential value, many CWR are threatened by habitat loss and degradation (Fischer & Lindenmayer, 2007; Kell et al., 2011), invasive species (Diaz et al., 2006; Ford-Lloyd et al., 2011) and climate change (Jarvis, Lane, et al., 2008). A variety of CWR conservation efforts are forming a response (Khoury, Greene, et al., 2019), both ex situ (in botanic gardens and seed banks) and in situ (in protected areas). A lack of representativeness of species and their intraspecific diversity has been recognized in genebanks (Castañeda-Álvarez et al., 2016) and in protected areas (Heywood et al., 2007; Khoury, Amariles, Soto, Díaz, Sotelo, Sosa, Ramirez-Villegas, Achicanoy, Velásquez-Tibatá, et al., 2019; Maxted et al., 2013).

1.2 | Domesticated sorghum

Here, we refer to domesticated sorghum as the many varieties of the species Sorghum bicolor (L.) Moench, including the cultivated varieties of the crop’s progenitor, S. bicolor subsp. verticilliflorum (Steud.) de Wet ex Wiersema & J. Dahlb. (also known as Sorghum arundinaceum [Desv.] Stapf). Sorghum was domesticated in eastern Africa around 5,000 years ago as a source of grain for human consumption (Fuller & Stevens, 2018). Today, it is grown on every inhabited continent and is the fifth-most important cereal crop globally in terms of tons produced (FAO, 2019). Its predominant use remains human consumption, especially as a grain in sub-Saharan Africa, with its ability to grow without fertilizer being advantageous in subsistence systems (Hadebe et al., 2017). There is also widespread use of sorghum in the production of syrup and alcoholic beverages, and a growing market for gluten-free products (Aruna & Visarada, 2019). In developed countries, its major use is as animal feed, with pigs and chickens fed on the grain and cattle fed on the stem and leaves (Ronda et al., 2019). Sorghum is also grown for bioethanol production, with yield per hectare generally equaling that of maize and exceeding it under dry conditions (Putnam et al., 1991). One of sorghum’s most notable agronomic traits is its superior drought and heat tolerance compared with other cereals (Dai, 2013; Hadebe et al., 2017; Rosenow & Clark, 1981).

Like many domesticated crops, sorghum exhibits genetic uniformity as a result of intensive selection for traits such as drought resistance and yield (Doebley et al., 2006). Sorghum diversification breeding with CWR has not advanced as far as in other major cereal crops, in part due to incompatibility constraints (Hodnett et al., 2005).

Fortunately, the introgression of traits from CWR into sorghum lines has recently become more achievable with the advent of S. bicolor with success rates of up to 20.7% (Liu & Godwin, 2012), compared with just 0.286% in the first published attempts (Casas et al., 1993). This progress potentially allows a greater use of wild Sorghum Moench (and other genera), which cannot be crossed with the crop using conventional techniques.

1.3 | Sorghum’s wild relatives

The genus Sorghum is currently considered to contain 22 wild taxa, whose collective range extends from Australia to the Pacific Islands, Southeast, East and South Asia, Central America and much of sub-Saharan Africa (Table 1). Seventeen wild taxa are native to Australia, with 13 being endemic, even though the crop itself was domesticated in Africa (Dillon, Shapter et al., 2007). Despite having a negligible contribution to the domestication of globally important crops (Khoury et al., 2016), Australia’s proximity to Asia and the Pacific Islands has engendered a surprising diversity of CWR, including those of sorghum, bananas and rice (Norton et al., 2017). The majority of Australian Sorghum taxa are located in the northern, monsoonal region of the country (Andrew & Mott, 1983; Lazarides et al., 1991), mainly occurring in the Northern Territory, Western Australia and Queensland. Sorghum trichocladum (Rupr. ex Hack.) Kuntze is the only species native to the Americas, with a distribution between southern Mexico and Honduras. The five remaining taxa are distributed across Africa and Asia, including the two taxa most closely related to domesticated sorghum—S. bicolor subsp. verticilliflorum and Sorghum propinquum (Kunth.) Hitchc.—which respectively have broad distributions across sub-Saharan Africa and eastern Asia.

Most wild Sorghum taxa are able to adapt to a range of edaphic conditions and collectively cover a broad range of habitats, including rocky slopes, sand dunes, grasslands and forests (Lazarides et al., 1991). This suggests that the CWR might contain high levels of genetic variation across populations.

Various traits of sorghum’s CWR have already been identified as potentially useful for introduction into S. bicolor, including resistance to pests such as sorghum shoot fly and spotted stem borer (Kamala et al., 2009; Venkateswaran, 2003), resistance to sorghum downy mildew (Kamala et al., 2002) and low cyanogenic glucoside concentrations (Cowen et al., 2020; Table 2). There is interest in expanding sorghum’s environmental tolerance, especially tolerance to colder climates (Fiedler et al., 2016; Yu & Tuinstra, 2001), which could potentially be achieved through the use of CWR from colder environments such as Sorghum leiocladum (Hack.) C. E. Hubb. Wild Sorghum taxa native to the Katherine Region in northern Australia...
| Taxon | Subgenus | Gene pool | Longevity | Native range | Conservation designation |
|-------|----------|-----------|-----------|--------------|--------------------------|
| Cleistachne sorghoides Benth. | N/A | 3 | Annual | Eastern Africa and India | N/A |
| Sorghum amplum | Stiposorghum | 3 | Annual | Western Australia | N/A |
| Sorghum angustum S. T. Blake | Stiposorghum | 3 | Annual | Northeast Queensland, Australia | RL—Least Concern |
| Sorghum bicolor (L.) Moench | Eusorghum | 1 | Annual | Africa | N/A |
| Sorghum brachypodum | Stiposorghum | 3 | Annual | Kakadu National Park, Australia | ALA—Least Concern |
| Sorghum bulbosum | Stiposorghum | 3 | Annual | Northern Australia | ALA—Least Concern |
| Sorghum ecarinatum | Stiposorghum | 3 | Annual | Northern Australia | ALA—Least Concern |
| Sorghum extans | Stiposorghum | 3 | Annual | Northern Territory, Australia | ALA—Least Concern |
| Sorghum grande | Parasorghum | 3 | Perennial | Northern Australia | ALA—data deficient |
| Sorghum interjectum | Stiposorghum | 3 | Annual/ Perennial | Northern Australia | RL—Least Concern |
| Sorghum intrans F. Muell. ex Benth. | Stiposorghum | 3 | Annual | Northern Territory, Australia | ALA—Least Concern |
| Sorghum laxiflorum F. M. Bailey | Heterosorghum | 3 | Annual | Northern Australia and Papua New Guinea | ALA—Least Concern |
| Sorghum leiocladum (Hack.) C. E. Hubb. | Parasorghum | 3 | Perennial | Eastern Australia | ALA—Least Concern (NT), Endangered (VIC) |
| Sorghum macrospernum E. D. Garber | Chaetosorghum | 3 | Annual | Northern Territory, Australia | ALA—Near Threatened |
| Sorghum matarankense E. D. Garber & Snyder | Parasorghum | 3 | Annual | Northern Territory, Australia | ALA—Least Concern |
| Sorghum nitidum (Vahl) Pers. | Parasorghum | 3 | Perennial | Queensland (Australia), New Guinea, Southeast Asia, and the Indian subcontinent | N/A |
| Sorghum plumosum (R. Br.) P. Beauv. | Stiposorghum | 3 | Perennial | Northern Australia and Indonesia | ALA—Least Concern |
| Sorghum propinquum (Kunth) Hitchc. | Eusorghum | 1 | Perennial | Southern India, Sri Lanka, southern China, Taiwan, and Southeast Asia | N/A |
| Sorghum purpureoserceum (Hochst. ex A. Rich.) Schweinf. & Asch. | Parasorghum | 3 | Annual | India, the Sahel, and east and west tropical Africa | RL—Least Concern |

(Continues)
have also historically been used as food sources by Dagoman people (Arndt, 1961), showing that they are already palatable and may even show promise as new crops themselves. Unfortunately, much about the life history and conservation status of sorghum’s wild relatives has not yet been documented (Ananda et al., 2020).

This study aims to provide a further understanding of the eco-geographic adaptations, distributions and conservation status of wild sorghums. To do this, we characterized the climatic and topographic niches of wild *Sorghum* taxa, calculated species distribution models using occurrence information combined with climatic and topographic data, used these models to assess the current conservation of wild *Sorghum* taxa both ex situ and in situ and conducted preliminary threat assessments for the taxa.

## METHODS

### 2.1 Study taxa

In this paper, we analysed the distribution and conservation status of all 22 known wild taxa of the genus *Sorghum* as listed by USDA ARS NPGS (2020; Table 1). Landraces of domesticated sorghum, referred to as “wild” by some authors (Mace et al., 2013), were not included. *Cleistachne sorghoides* Benth. was included as part of the genus in this study due to molecular evidence placing it within the *Sorghum* clade (Dillon, Lawrence, et al., 2007; Liu et al., 2014; Sun et al., 1994), despite its nomenclature not yet reflecting this evidence (Table 1).

The study included wild individuals of *S. bicolor* subsp. *verticilliflorum*, but not the domesticated subspecies of *S. bicolor* or *S. bicolor* hybrids: *Sorghum halepense* (L.) Pers., *Sorghum × almum* Parodi and *S. bicolor* nothosubsp. *drummondii* (Steu.) de Wet ex Davidse, all of which were produced through hybridization of domesticated sorghum with wild taxa. *Sorghum halepense* is commonly found beyond its native range and is considered a noxious weed in many regions (Holm et al., 1977), compounding its lack of suitability for this study.

*Sorghum bicolor* subsp. *verticilliflorum* and *S. propinquum* are considered part of sorghum’s primary gene pool, with all other taxa being in the tertiary gene pool (Harlan & de Wet, 1971; USDA ARS NPGS, 2020; Table 1). Taxonomic names were standardized as per USDA ARS NPGS (2020).

### 2.2 Occurrence data

Occurrence data were compiled from the Global Biodiversity Information Facility (GBIF; GBIF, 2019), the Smithsonian Collections (Smithsonian, 2020) and the Crop Wild Relative Occurrence
Database (Global Crop Diversity Trust, 2019a). Ex situ conservation occurrence data were compiled from the Genesys Plant Genetic Resources portal (Genesys-PGR; Global Crop Diversity Trust, 2019b), from the GRIN-Global portal of the USDA National Plant Germplasm System (GRIN-GLOBAL; USDA ARS NPGS, 2020) and through direct communication with the Australian Grains Genebank. Duplicates in the databases, as well as records listed as grown in institutes, farms or home gardens, were removed.

Records were then either classified as “G” (for ex situ records sourced primarily from genebanks) or “H” (for reference records sourced mostly from herbaria). For GBIF data, all records listed as “living specimen” were labelled G, with other categories (observation, literature, preserved specimen, human observation, machine observation, material sample and unknown) labelled H. All records from Genesys-PGR and the Australian Grains Genebank were labelled G. Active and inactive records in GRIN-Global were labelled G and H, respectively. G occurrences that had detailed locality information but no coordinates were georeferenced by hand in Google Maps (Google, 2019) in order to maximize the completeness of the G dataset. Occurrence records were then uploaded to ArcMap version 10.6.1 (Esri, 2018), with occurrences either corrected or removed if they were located in water or clearly incorrect locations. The final occurrence dataset is available in Appendix S1 in Supporting Information.

2.3 | Species distribution modelling

Species distribution models were created using the maximum entropy (MaxEnt) algorithm (Phillips et al., 2006, 2017) in the R package “dismo” (Hijmans, Phillips, et al., 2017). Following Khoury, Amariles, Soto, Diaz, Sotelo, Sosa, Ramírez-Villegas, Achicanoy, Velásquez-Tibatá, et al. (2019), models were produced using 26 eco-geographic variables (Table S2.1 in Supporting Information), including 19 bioclimatic variables, solar radiation, water vapour pressure and wind speed, all of which were derived from WorldClim 2.0 (Fick & Hijmans, 2017). For the final three variables, we produced annual values by calculating the median across monthly values. We also included altitude, which was compiled from the CGIAR-CSI dataset based on NASA Shuttle Radar Topography Mission data (Jarvis, Reuter, et al., 2008); and slope and aspect, which were calculated from the altitude data using the terrain function in the R package “raster” (Hijmans, 2017). All eco-geographic variables were processed at a 2.5-arc-minute spatial resolution (approximately 5 km² at the equator). The eco-geographic variables used in MaxEnt models were selected separately for each taxon using the R package “VSURF” (Genuer et al., 2019). Variables were ranked in order of impact on model performance, and every variable that made no measurable impact was removed. The remaining variables were tested for Pearson’s correlation with other variables, and any variable with a correlation coefficient greater than 0.7 or less than −0.7 with any variable more important than itself was removed. This process was repeated until there were no pairs of variables within the five most important variables with a coefficient greater than 0.7 or less than −0.7.

For each taxon, a spatial background was created based on the boundaries of the ecoregions in which taxon occurrences were located (Olson et al., 2001). Pseudoabsence numbers were then created in proportion to the area of the taxon’s spatial background, with a maximum of 5,000 pseudoabsences. Ten replicate models were produced for each taxon using the MaxEnt algorithm (K = 10), using linear, quadratic, hinge and product features, with a regularization parameter β = 1.0. The median of these replicates formed the final MaxEnt model. Median models were evaluated using three measures: area under the receiver operating characteristic curve (AUC); standard deviation of the AUC across replicates (SDAUC); and the proportion of the potential distribution model with a standard deviation above 0.15 (ASD15). For a model to be considered accurate, each of the following criteria must be met: AUC ≥ 0.7; SDAUC < 0.15; and ASD15 ≤ 10%. Lastly, MaxEnt models were thresholded using the maximum sum of sensitivity and specificity (Liu et al., 2005, 2013). Models were clipped as required to the extent of the taxon’s spatial background.

2.4 | Ecogeographic characterization

Ecogeographic predictor data, at a resolution of 2.5 arc minutes for the 26 ecogeographic variables from the WorldClim 2.0 and CGIAR-CSI datasets, were extracted for all georeferenced records for all taxa (Appendix S1 in Supporting Information). These data were used to characterize taxa in terms of their potential ecogeographic niches for each variable. We also assessed the representation of these niches in ex situ conservation by comparing the distributions of a taxon’s G occurrences within its full spread of occurrences.

2.5 | Conservation gap analysis

The ex situ and in situ conservation of each taxon was assessed following Khoury, Carver, Barchenger, et al. (2020) and Khoury, Carver, Kates, et al. (2020), with four scores calculated for both ex situ and in situ schemes. All scores had a scale of 0–100, with 0 representing extremely poor conservation, and 100 representing complete conservation.

The first ex situ score was the sampling representativeness score \( \text{SRS}_{\text{ex}} \), which is the ratio of G occurrences to H occurrences. Unlike the other scores, \( \text{SRS}_{\text{ex}} \) takes into account both georeferenced and non-georeferenced G occurrences. The second ex situ score was the geographic representativeness score \( \text{GRS}_{\text{ex}} \). To calculate this score, 50-km-radius buffers were created around each G occurrence. \( \text{GRS}_{\text{ex}} \) is the percentage of the taxon’s thresholded distribution model that is covered by these G occurrence buffers. The third ex situ score was the ecological representativeness score \( \text{ERS}_{\text{ex}} \). This score made use of a raster layer, which divides the terrestrial world into 867 ecoregions (Olson et al., 2001), as well as the buffers around G occurrences. \( \text{ERS}_{\text{ex}} \)
is the percentage of ecoregions included in the taxon’s distribution model that feature at least once in the taxon’s G occurrence buffers.

The first in situ score was the sampling representativeness score (SRS), which is the percentage of total occurrences that lie inside the protected areas marked as “designated,” “inscribed” or “established” in the World Database of Protected Areas (IUCN, 2019a). The second in situ score was the geographic representativeness score (GRS), which is the percentage of the taxon’s entire thresholded distribution model of that overlaps with the protected area layer. The third in situ score was the ecological representativeness score (ERS), which is the percentage of ecoregions included in the taxon’s thresholded distribution model that are featured in the overlap between the thresholded model and protected area layer.

Final scores for ex situ (FCS_ex) and in situ (FCS_in) conservation were calculated by taking the mean of the three ex situ scores and the mean of the three in situ scores, respectively. A final combined conservation score for the taxon (FCS_c) was then calculated by taking the mean of FCS_ex and FCS_in. All FCS scores had a scale of 0–100 and were further categorized where FCS < 25 signifies a high-conservation-priority (HP) taxon, 25 ≤ FCS ≤ 50 medium-priority (MP) taxon and 50 ≤ FCS ≤ 75 low-priority (LP) taxon, and 75 ≤ FCS ≤ 100 signifies that the taxon may be considered sufficiently conserved (SC).

2.6 | Preliminary threat assessment

To complement the conservation gap analysis, we also used the occurrence datasets to calculate two metrics adapted from the IUCN Red List criteria (IUCN Standards & Petitions Committee, 2019): the extent of occurrence (EOO) and area of occupancy (AOO). EOO estimates how widespread a taxon is by creating the shortest possible continuous boundary encompassing all occurrence points and calculating the area inside the boundary. AOO estimates the area within the EOO, which is actually occupied by a taxon by calculating the minimum number of 2 km × 2 km grids required to cover all occurrence points. These calculations were performed using the R package “redlistr” (Lee et al., 2019). Taxa were categorized using both metrics, whereby a taxon is Critically Endangered when EOO < 100 km² or AOO < 10 km²; Endangered when 100 km² < EOO < 5,000 km² or 10 km² < AOO < 500 km²; Vulnerable when 5,000 km² < EOO < 20,000 km² or 500 km² < AOO < 2,000 km²; Near Threatened when 20,000 km² > EOO < 45,000 km² or 2,000 km² < AOO < 4,500 km²; and Least Concern when EOO ≥ 45,000 km² and AOO ≥ 4,500 km² (IUCN Standards & Petitions Committee, 2019). While these metrics do not provide the full set of criteria needed for classification on the Red List, they offer indications of the threat status of each taxon.

3 | RESULTS

A total of 13,846 H records and 654 G records (of which 540 had coordinates) were compiled for analysis, with taxon occurrence numbers ranging from 40 records for S. macrospernum to 4,208 records for Sorghum plumosum (R. Br.) P. Beauv. All taxa had adequate occurrences for distribution modelling (van Prooijen et al., 2016), and all models passed the evaluation criteria (Table S2.3 in Supporting Information).

3.1 | Taxon distributions

The predicted range of wild Sorghum includes eastern and northern Australia, South, Southeast and East Asia, Papua New Guinea, Central America and much of sub-Saharan Africa (Figure 1). The primary gene pool, despite consisting of just two taxa, covers much of this range, with S. bicolor subsp. verticilliflorum distributed across much of sub-Saharan Africa and S. propinquum extending from southern China to New Guinea (Appendix S3 in Supporting Information). The tertiary gene pool is also extremely widespread, with taxon density at its greatest in northern Australia, peaking at 11 taxa in some 5 km² cells in the Katherine Region, Northern Territory. Other regions contain much lower levels of tertiary gene pool taxon richness, including Central America with one taxon (S. trichocladum) and Africa with just three taxa (C. sorghoides, Sorghum purpureoserratum [Hochst. ex A. Rich.] Schweinf. & Asch. and Sorghum versicolor Andersson).

The most widespread taxon is Sorghum nitidum (Vahl) Pers., whose distribution model covers much of eastern Asia, ranging latitudinally from central Japan to southeast Australia and extending as far west as Pakistan. This was reflected in S. nitidum having the largest EOO of the tested taxa (34,403,804 km²). The least widespread taxon is S. macrosporum, which is limited to a small section of the Katherine Region, and has an EOO of just 400 km².

3.2 | Ecogeographic characterization

Regarding ecogeographic niches, substantial variation was found between taxa. The CWR able to survive in the most extreme climatic niches, measured by median of occurrences, included the following: Sorghum bulbosum Lazarides, Sorghum matarankense E. D. Garber & Snyder and Sorghum timorense (Kunth) Büse with the highest maximum temperatures in the warmest month of the year; S. leiocladum, S. nitidum and S. versicolor with the lowest minimum temperatures in the coldest month; Sorghum amplum Lazarides and Sorghum extans Lazarides with the highest precipitation in the wettest month; and all taxa but Sorghum angustum S. T. Blake and S. propinquum having occurrences in sites with no precipitation in the driest month. Sorghum leiocladum was also of note due to its median value for annual mean temperature of 15.19°C, the only taxon with a value below 20°C and with just three other taxa having medians below 25°C (Figures S2.1 and S2.2 in Supporting Information).

3.3 | Conservation gap analysis

The majority of sorghum taxa (19 out of 23) were determined to be medium priorities overall for further conservation action, with
**FIGURE 1** Predicted taxonomic richness map combining the 23 wild Sorghum Moench species distribution models in (a) the genus’ entire native range and (b) northern Australia. Warmer colours indicate greater numbers of taxa potentially coexisting in the same areas.

**FIGURE 2** Conservation gap analysis results per Sorghum Moench taxon. Taxa are listed by descending priority for further conservation action by priority categories (high priority [HP] [red]; medium priority [MP] [orange]; low priority [LP] [yellow]; and sufficiently conserved [SC] [green]). Results of the conservation assessments within each strategy (sampling representativeness score ex situ [SRS\textsubscript{ex}]; in situ [SRS\textsubscript{in}]; geographic representativeness score ex situ [GRS\textsubscript{ex}]; in situ [GRS\textsubscript{in}]; and ecological representativeness score ex situ [ERS\textsubscript{ex}]; in situ [ERS\textsubscript{in}]) are also displayed.
### Table 3 Conservation gap analysis results per wild Sorghum Moench taxon

| Taxon                                      | Total records | Total G records | SRS<sub>ex</sub> | GRS<sub>ex</sub> | ERS<sub>ex</sub> | FCS<sub>ex</sub> | SRS<sub>in</sub> | GRS<sub>in</sub> | ERS<sub>in</sub> | FCS<sub>in</sub> | FCSc | Priority category | Red List category recommendation |
|--------------------------------------------|---------------|-----------------|------------------|------------------|------------------|------------------|-----------------|-----------------|------------------|------------------|------|-------------------|-------------------------------|
| Cleistachne sorghoides                     | 66            | 6               | 10.00            | 1.84             | 16.67            | 9.50             | 14.29           | 19.91           | 95.83            | 43.34            | 33.69 | MP                | EN                           |
| Sorghum amplum                             | 43            | 2               | 4.88             | 6.88             | 50.00            | 20.59            | 51.16           | 47.90           | 100.00           | 66.35            | 43.47 | MP                | EN                           |
| Sorghum angustum                           | 142           | 11              | 8.40             | 40.98            | 60.00            | 36.46            | 19.01           | 32.22           | 100.00           | 50.41            | 43.43 | MP                | EN                           |
| Sorghum bicolor subsp. verticilliflorum    | 857           | 180             | 26.59            | 5.33             | 23.44            | 18.45            | 13.15           | 17.85           | 87.50            | 39.50            | 28.98 | MP                | NT                           |
| Sorghum brachypodium                       | 139           | 3               | 2.21             | 20.20            | 50.00            | 24.14            | 82.01           | 100.00          | 66.35            | 43.47            | MP    | EN                | LP                           |
| Sorghum bulbosum                           | 288           | 22              | 8.27             | 46.82            | 75.00            | 43.36            | 7.64            | 23.27           | 100.00           | 43.64            | 43.5  | MP                | VU                           |
| Sorghum ecarinatum                         | 59            | 12              | 25.53            | 48.56            | 75.00            | 49.70            | 25.42           | 11.72           | 100.00           | 45.72            | 47.71 | MP                | EN                           |
| Sorghum extans                             | 74            | 4               | 5.71             | 30.13            | 100.00           | 45.28            | 10.81           | 35.85           | 100.00           | 48.89            | 47.09 | MP                | EN                           |
| Sorghum grande                             | 42            | 3               | 7.69             | 9.17             | 50.00            | 22.29            | 14.29           | 31.21           | 100.00           | 48.50            | 35.39 | MP                | EN                           |
| Sorghum interjectum                        | 220           | 20              | 10.00            | 14.76            | 80.00            | 34.92            | 26.36           | 29.85           | 100.00           | 50.41            | 43.5  | MP                | VU                           |
| Sorghum intrans                            | 1927          | 40              | 2.12             | 30.79            | 66.67            | 33.19            | 43.63           | 35.85           | 100.00           | 59.83            | 46.51 | MP                | NT                           |
| Sorghum laxiflorum                         | 391           | 26              | 7.12             | 21.50            | 58.33            | 28.98            | 22.11           | 24.75           | 75.00            | 40.62            | 34.8  | MP                | VU                           |
| Sorghum leiocladum                         | 2,679         | 36              | 1.36             | 27.14            | 71.43            | 33.31            | 22.36           | 18.38           | 100.00           | 46.91            | 40.11 | MP                | LC                           |
| Sorghum macropernum                        | 40            | 4               | 11.11            | 80.14            | 100.00           | 63.75            | 0.00            | 0.00            | 0.00             | 0.00             | 31.88 | MP                | EN                           |
| Sorghum mataanense                         | 80            | 11              | 15.94            | 15.46            | 50.00            | 27.13            | 11.54           | 13.11           | 100.00           | 41.55            | 34.34 | MP                | EN                           |
| Sorghum nitidum                            | 506           | 18              | 3.69             | 4.31             | 10.87            | 6.29             | 11.53           | 12.88           | 71.74            | 32.05            | 19.17 | HP                | LC                           |
| Sorghum plumosum                           | 4,208         | 67              | 1.62             | 25.29            | 64.29            | 30.40            | 32.94           | 25.50           | 85.71            | 48.05            | 39.22 | MP                | LC                           |
| Sorghum propinquum                         | 42            | 7               | 20.00            | 2.41             | 4.55             | 8.98             | 8.11            | 16.43           | 90.91            | 38.48            | 23.73 | MP                | EN                           |
| Sorghum purpureosericum                    | 162           | 67              | 71.28            | 11.26            | 52.17            | 44.90            | 16.00           | 19.11           | 82.61            | 39.24            | 42.07 | MP                | EN                           |
| Sorghum stipoidum                          | 1,013         | 53              | 5.52             | 20.90            | 80.00            | 35.47            | 22.41           | 28.04           | 100.00           | 50.15            | 42.81 | MP                | NT                           |
| Sorghum timorense                          | 1,279         | 41              | 3.31             | 22.37            | 55.56            | 27.08            | 18.42           | 18.03           | 88.89            | 41.78            | 34.43 | MP                | NT                           |
| Sorghum trichocladum                       | 52            | 0.00            | 0.00             | 0.00             | 0.00             | 0.00             | 11.54           | 11.31           | 90.91            | 37.92            | 25.56 | MP                | EN                           |
| Sorghum versicolor                         | 192           | 21              | 12.88            | 4.11             | 43.33            | 19.91           | 25.56           | 20.93           | 83.33            | 43.27            | 31.59 | MP                | VU                           |

*Note: Results of the conservation assessments within each strategy (sampling representativeness score ex situ [SRS<sub>ex</sub>], geographic representativeness score ex situ [GRS<sub>ex</sub>], and ecological representativeness score ex situ [ERS<sub>ex</sub>]; and sampling representativeness score in situ [SRS<sub>in</sub>], geographic representativeness score in situ [GRS<sub>in</sub>], and ecological representativeness score in situ [ERS<sub>in</sub>]) are displayed. The final conservation score combined (FCSc) is the average of the final ex situ (FCS<sub>ex</sub>) and in situ (FCS<sub>in</sub>) scores. The FCSc is used to categorize taxa for further conservation action: high priority (HP, FCSc < 25), medium priority (MP, 25 ≤ FCSc < 50), low priority (LP, 50 ≤ FCSc < 75), and sufficiently conserved (SC, FCSc ≥ 75). Red List category recommendations are based on extent of occurrence (EOO) and area of occupancy (AAO). Red List categories: Critically Endangered (CR), Endangered (EN), Vulnerable (VU), Near Threatened (NT) and Least Concern (LC).*
three taxa being high priorities (S. nitidum, S. propinquum and S. trichocladum), and just one taxon low priority (Sorghum brachypodum Lazarides). FCSc results ranged from 18.96 to 51.10 (Figure 2; Table S2.2 in Supporting Information).

The range in comprehensiveness of ex situ conservation was greater than that of FCSc, with FCS\textsubscript{ex} varying from 0 to 63.75. Nine taxa were classified as high priorities for ex situ conservation, with 13 taxa being of medium priority, and one (S. macrospermum) low priority (Figure 2).

The range in comprehensiveness of in situ conservation was also greater than that of FCSc, with FCS\textsubscript{in} ranging from 0 to 78.07. Within taxa, FCS\textsubscript{in} values were generally greater (i.e. indicating a better current state of conservation) than corresponding FCS\textsubscript{ex} values, with S. macrospermum (whose FCS\textsubscript{in} was 0) being the only taxon not following this trend. Only S. macrospermum was classified as a high priority for in situ conservation, with 16 taxa being classified as medium priorities, five as low priorities and one (S. brachypodum) as sufficiently conserved (Figure 2).

The preliminary threat assessment indicated that S. leiocladum and S. pluminosum may be considered to be listed as Least Concern; S. bicolor subsp. verticilliflorum, Sorghum intrans F. Muell. ex Benth., Sorghum stipoidum (Ewart & Jean White) C. A. Gardner & C. E. Hubb. and S. timorenses as Near Threatened; S. bulbosum, Sorghum interjectum Lazarides, Sorghum laxiflorum F. M. Bailey, S. nitidum and S. versicolor as Vulnerable; and the remaining 12 taxa as Endangered (Table 3).

4 | DISCUSSION

4.1 | General patterns

As has been found in other clades (Khoury, Carver, Barchenger, et al., 2020; Lebeda et al., 2019), Sorghum’s in situ conservation scores were generally higher than corresponding ex situ scores. This indicates the potential value of in situ conservation to CWR protection, with the possibility for many taxa to be protected by a single well-placed protected area (Maxted et al., 2013). This is, of course, subject to field verification of taxon presences and sound protected area management (Svancara et al., 2005). Despite occurring within a protected area, without monitoring and management plans a taxon is Vulnerable (Mason et al., 2015; Pressey et al., 2015).

Predictably, national genebanks (with the exception of the Millennium Seed Bank) primarily store germplasm of wild Sorghum taxa native to their own regions. For example, the Australian Grains Genebank does not have more than seven different accessions of any Sorghum taxon that is not native to Australia, even in the case of S. bicolor subsp. verticilliflorum. While this trend is understandable, increased sharing of germplasm between genebanks, while avoiding excessive duplication, could aid in increasing the efficiency with which it can be distributed to local crop developers and researchers in each region, maximizing the genetic diversity available to them.

Although gap analysis scores were calculated across taxa in a consistent manner, potential spatial biases in the underlying datasets (Beck et al., 2014) could have affected distribution models and therefore taxon gap analysis scores to varying degrees. To mitigate this challenge, taxa are separated based on native region in the remainder of the discussion. Species native to Africa and Asia were lumped due to multiple taxa having distributions in both continents.

The preliminary threat assessments of thirteen taxa did not match their current Red List determination (Table 3). While this might potentially suggest a need for revisions of the categorization of these taxa, our primary assessment did not include additional steps, including change over time analyses and expert discussion, which are incorporated into official Red List assessments. Our assessments were solely based on EOO and AOO, with AOO determining the overall categorization for every Sorghum taxon (Table S2.4 in Supporting Information), despite AOO’s potential to greatly underestimate true range size (Sheh et al., 2012).

4.2 | African and Asian taxa

Of the taxa native to Africa and Asia, only S. propinquum was classified a high-priority taxon for further conservation, with every other taxon being medium priorities. Sorghum propinquum’s relatively low ex situ conservation score for the region (8.98) is especially concerning due to this species being in domesticated sorghum’s primary gene pool. The other taxon in sorghum’s primary gene pool, S. bicolor subsp. verticilliflorum, had a slightly higher ex situ conservation score (18.45), which potentially reflects the taxon’s great historical use in sorghum breeding as the crop’s progenitor, but again highlights room for improvement in the protection of this taxon.

4.3 | Australian taxa

No endemic Australian Sorghum taxa were listed as high priorities for further conservation overall. There is, however, room for improvement. Three of these taxa were considered high priorities for further ex situ conservation: S. amplum, Sorghum grande Lazarides and, despite it having the best FCSc score in the genus, S. brachypodum. These low ex situ scores are unsurprising considering that these taxa had just two, three and three G accessions, respectively. Further seed collection efforts would improve protections of these taxa, with the geographic priorities for further collection of S. grande and S. brachypodum centred in the Katherine Region. This region has the greatest predicted richness of Sorghum taxa in the world (Figure 1), meaning further seed collection there could be an efficient means by which to improve ex situ conservation across the tertiary gene pool. In terms of priorities for in situ conservation, S. macrospermum was the only taxon with a high-priority FCS\textsubscript{ex} as none of its narrow distribution lie within a protected area. This is worrying, especially with its EOO corresponding to a potential Endangered Red List status. However, its low-priority FCS\textsubscript{ex} score somewhat mitigates this
issue, with much of its estimated intraspecific genetic diversity already safeguarded in genebanks. These results generally support the recommendations of Norton et al. (2017) for enhanced ex situ conservation of wild Australian *Sorghum* taxa.

Non-endemic *Sorghum* native to Australia (*S. laxiflorum, S. nitidum, S. plumosum* and *S. timorense*) were generally classified as medium priorities for further conservation, with the exception of *S. nitidum*, a high-priority species. One possible reason for *S. nitidum’s* low FCS\textsubscript{ex} score (6.29) is the concentration of reported G accessions coming from Australia—eight out of the 14 georeferenced G accessions for the taxon come from the country, despite the taxon also having occurrences in ten other nations. It is unclear whether seeds have not been collected from these countries, or whether collections have simply not been reported on openly accessible databases. Areas such as the Ryukyu Islands and coastal Taiwan could be good targets for future seed collections due to their accessibility and relatively small areas (compared with rural areas of China and Papua New Guinea, for example).

Future attempts to increase in situ protection for Australian *Sorghum* species would be most efficient in the taxon-rich Katherine Region. *Sorghum macrospernum* and *S. nitidum*—the taxa with the two lowest FCS\textsubscript{in} scores of the Australian *Sorghum* species—are both native to this region.

### 4.4 Central American taxa

*Sorghum trichocladum* is the only taxon in the genus native to the Americas, as well as being the only one currently without any ex situ germplasm accessions documented on openly accessible platforms. It is currently unclear whether there is indeed no germplasm available for this species in genebanks, or whether collections have not yet been identified or reported. Fortunately, *S. trichocladum’s* distribution model significantly overlaps current protected areas, though field verification is needed to confirm these distributions. Its ERS\textsubscript{in} score of 90.91 is particularly positive and suggests that in situ protections may be well distributed across the different ecoregions in which the species is found.

### 4.5 Challenges and limitations

There exist several limitations regarding the calculation and use of species distribution models, which should be acknowledged when considering the results of this study. Firstly, there are inevitable gaps in occurrence datasets for taxa that have not been fully sampled. This can lead to the exclusion of some areas of actual ranges in distribution models if these areas are not represented in available datasets. Also, gaps in georeferencing data also could have affected model accuracy, as well as influencing conservation scores. Secondly, spatial bias towards roadsides and other areas of human activity can impact models built from presence-only data (Hijmans, 2012). These issues are commonplace when using openly available occurrence datasets, but we attempted to mitigate them by producing ten replicate models for each taxon using different random splits between testing and training data. Models were also made more conservative by limiting taxon backgrounds to ecoregions in which taxon occurrence data existed. As mentioned, spatial biases and data availability issues generally affect data from developing countries more than developed ones, potentially leading to inconsistencies in the accuracy of underlying data from the different regions *Sorghum* is native to (Beck et al., 2014). A final limitation is that our models took 26 eco-geographic variables into account, but did not include some other factors that influence taxon distributions, including biotic interactions, edaphic variables and recent habitat degradation. The 2.5-arc-minute spatial resolution used can also lead to some microclimatic conditions within grid cells being overlooked, as well as models potentially being too general in their determinations of "presence" of a species within an environmentally heterogeneous cell. For these reasons, our distribution models should be considered planning tools to guide explorations for confirmation in the field, and not definitive guides of where a taxon is and is not present.

Additionally, there has been debate over the monophyletic status of *Sorghum*, as well as over which species belong in this genus (e.g. Dillon, Lawrence, et al., 2007; Hawkins et al., 2015; Kellogg, 2013; Spangler, 2003). For this reason, our knowledge about these CWR, in terms of conservation and use in crop improvement, should continue to be updated according to the most recent classifications of *Sorghum* taxa. Readers should also be aware of changing data regarding the distributions of taxa (whether they are extirpated in an area, or found in previously unknown areas) and consider our study in the light of these developments.

### 4.6 Future directions

Due to the presence of *S. bicolor* subsp. *vernicilliflorum* and *S. propinquum*, from sorghum’s primary gene pool, in Africa and Asia, respectively, it is particularly important that further ex situ *Sorghum* conservation efforts are made in these regions. It remains unclear whether these taxa’s respective categorizations as medium and high priorities for further conservation are mainly due to uncomprensive efforts or online datasets not being as comprehensive as they are elsewhere. Regardless, improvement of the germplasm collections available to breeders and researchers is vital to current crop improvement efforts. Due to the immense difficulty of producing comprehensive, unbiased occurrence datasets across the genus, conservationists must act urgently, using knowledge already available, in order to ensure the persistence of CWR and their intraspecific diversity before populations decline further. The loss of CWR through extinction and extirpation is a constant threat with irreversible consequences. Further delay of conservation action to prevent these outcomes would be unwise.

In addition to further ex situ and in situ conservation efforts, there remain various actions that could be taken to maximize the value of sorghum’s wild relatives to agriculture. Firstly, although
improvements to the breeding process and genetic modification in sorghum have recently occurred, continued advances in the gene introgression process and increased acceptance of genetically modified crops by regulatory bodies and the public would help to maximize the simplicity of the process, and would consequently allow more widespread use of wild Sorghum (and any other) taxon in crop development. Increased distribution of knowledge and resources (such as S. bicolor plants that allow cross-species hybridization and ex situ germplasm accessions from different regions) between researchers, crop developers and farmers would also allow faster progress in sorghum improvement. Current knowledge about the general biology of these CWR is limited (Table 2). It is vital that further research, particularly on the physiology of sorghum's CWR and their responses to environmental conditions, is conducted in order to allow a better understanding of which CWR might be useful in sorghum improvement.

ACKNOWLEDGEMENTS

The authors thank the botanists, taxonomists, plant collectors, geospatial scientists and genetic resource professionals who compiled and made available the occurrence and ecogeographic information used in this analysis. This research was supported by the ARC Discovery Project DP180101011 to RJH and RG and the International Center for Tropical Agriculture. HM was supported by a Monash University Research Training Stipend and a Tim Healey Memorial Scholarship from the AW Howard Memorial Trust.

PEER REVIEW

The peer review history for this article is available at https://pubon ns.com/publ on/10.1111/ddi.13166.

DATA AVAILABILITY STATEMENT

Occurrence data, processed ecogeographic data, and interactive taxon-level modelling and conservation status results and metrics are provided in the Supporting Information. Associated ecogeographic and spatial input data are available through open access repositories (Khoury, Amariles, Soto, Díaz, Sotelo, Sosa, Ramirez-Villegas, Achicanoy, Castañeda-Álvarez, et al., 2019). All code implemented in the analysis is available at: https://github.com/dcarver1/cwrSDM.

ORCID

Harry Myrans https://orcid.org/0000-0003-2690-6188
Colin K. Khoury https://orcid.org/0000-0001-7893-5744

REFERENCES

Ananda, G. K. S., Myrans, H., Norton, S. L., Gleadow, R., Furtado, A., & Henry, R. J. (2020). Wild sorghum as a promising resource for crop improvement. Frontiers in Plant Science, 11(1108). https://doi.org/10.3389/fpls.2020.01108
Andrew, M. H., & Mott, J. J. (1983). Annuals with transient seed banks: The population biology of indigenous Sorghum species of tropical north-west Australia. Australian Journal of Ecology, 8(3), 265–276. https://doi.org/10.1111/j.1442-9993.1983.tb01324.x
Arndt, W. (1961). Indigenous Sorghum as food and in myth: The Tagaman Tribe. Oceania, 32(2), 109–112. https://doi.org/10.1002/j.1834-4461.1961.tb01745.x
Aruna, C., & Visarada, K. (2019). Sorghum grain in food and brewing industry. In C. Aruna, K. Visarada, B. V. Bhat, & V. A. Tonapi (Eds.), Breeding Sorghum for diverse end uses (pp. 209–228). Cambridge: Woodhead Publishing.
Atlas of Living Australia. (2019). Atlas of Living Australia.
Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species’ geographic distributions. Ecological Informatics, 19, 10–15. https://doi.org/10.1016/j.ecoinf.2013.11.002
Casas, A. M., Kononowicz, A. K., Zehr, U. B., Tomes, D. T., Axtell, J. D., Butler, L. G., Bressan, R. A., & Hasegawa, P. M. (1993). Transgenic sorghum plants via microprojectile bombardment. Proceedings of the National Academy of Sciences of the United States of America, 90(23), 11212–11216. https://doi.org/10.1073/pnas.90.23.11212
Castañeda-Álvarez, N. P., Khoury, C. K., Achicanoy, H. A., Bernau, V., Dempewolf, H., Eastwood, R. J., Guarino, L., Harker, R. H., Jarvis, A., Maxted, N., Müller, J. V., Ramirez-Villegas, J., Sosa, C. C., Struik, P. C., Vincent, H., & Toll, J. (2016). Global conservation priorities for crop wild relatives. Nature Plants, 2(4), 16022. https://doi.org/10.1038/nplants.2016.22
CBD. (2010). Global Strategy for Plant Conservation: The targets 2011–2020. Paper presented at the Conference of the Parties to the Convention on Biological Diversity, Tenth Meeting, 18–29 October, Nagoya.
CBD. (2019). Aichi Biodiversity Targets. Retrieved from https://www.cbd.int/sp/targets/
CBD. (2020). Preparations for the Post-2020 Biodiversity Framework. Retrieved from https://www.cbd.int/conferences/post2020
Cowan, M. F., Blomstedt, C. K., Norton, S. L., Henry, R. J., Møller, B. L., & Gleadow, R. (2020). Crop wild relatives as a genetic resource for generating low-cyanide, drought-tolerant Sorghum. Environmental and Experimental Botany, 169, 103884. https://doi.org/10.1016/j.envexpbot.2019.103884
Dai, A. (2013). Increasing drought under global warming in observations and models. Nature Climate Change, 3(1), 52. https://doi.org/10.1038/nclimate1633
Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C., & Guarino, L. (2017). Past and future use of wild relatives in crop breeding. Crop Science, 57(3), 1070–1082. https://doi.org/10.2136/crops ci2016.10.0885
Diaz, S., Fargione, J., Chapin, F. S. III, & Tilman, D. (2006). Biodiversity loss threatens human well-being. PLoS Biology, 4(8), e277. https://doi.org/10.1371/journal.pbio.0040277
Dillon, S. L., Lawrence, P. K., Henry, R. J., & Price, H. J. (2007). Sorghum resolved as a distinct genus based on combined ITS1, ndhF and Adh1 analyses. Plant Systematics and Evolution, 268(1–4), 29–43. https://doi.org/10.1007/s00660-007-0571-9
Dillon, S. L., Shapter, F. M., Henry, R. J., Cordeiro, G., Izquierdo, L., & Lee, L. S. (2007). Domestication to crop improvement: Genetic resources for Sorghum and Saccharum (Andropogoneae). Annals of Botany, 100(5), 975–989. https://doi.org/10.1093/aob/mcm192
Doebley, J. F., Gaut, B. S., & Smith, B. D. (2006). The molecular genetics of crop domestication. Cell, 127(7), 1309–1321. https://doi.org/10.1016/j.cell.2006.12.006
Dweikat, I. (2005). A diploid, interspecific, fertile hybrid from cultivated sorghum, Sorghum bicolor, and the common Johnsongrass weed Sorghum halepense. Molecular Breeding, 16(2), 93–101. https://doi.org/10.1007/s11032-005-5021-1
**BIOSKETCH**

Harry Myrans is a PhD student at Monash University studying sorghum's crop wild relatives, with a particular interest in their physiology and how they could be used to mitigate problems related to cyanogenesis in forage sorghum. The co-authors on this article comprise biogeographers, genetic resource professionals and plant physiologists.

Author contributions: H.M., C.K.K. and R.G. conceived the research; H.M. compiled and processed occurrence data; M.V.D., C.K.K. and D.C. developed the methodologies and compiled and processed ecogeographic predictor and conservation analysis data; H.M. and M.V.D. performed the analyses; H.M. analysed the results and wrote the article; and all authors edited the article and approved the manuscript before submission.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Myrans H, Diaz MV, Khoury CK, Carver D, Henry RJ, Gleadow R. Modelled distributions and conservation priorities of wild sorghums (*Sorghum* Moench). *Divers Distrib.* 2020;26:1727-1740. [https://doi.org/10.1111/ddi.13166](https://doi.org/10.1111/ddi.13166)