



OPEN ACCESS

EDITED BY

Andrés J. Cortés,
Colombian Corporation for
Agricultural Research (AGROSAVIA),
Colombia

REVIEWED BY

Paul Gepts,
University of California, Davis,
United States
Thomas M. Davis,
University of New Hampshire,
United States
Robert Philipp Wagensommer,
Free University of Bozen-Bolzano, Italy

*CORRESPONDENCE

Jacqueline Batley
Jacqueline.batley@uwa.edu.au

SPECIALTY SECTION

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

RECEIVED 01 August 2022

ACCEPTED 25 October 2022

PUBLISHED 17 November 2022

CITATION

Tirnaz S, Zandberg J, Thomas WJW,
Marsh J, Edwards D and Batley J
(2022) Application of crop wild
relatives in modern breeding: An
overview of resources, experimental
and computational methodologies.
Front. Plant Sci. 13:1008904.
doi: 10.3389/fpls.2022.1008904

COPYRIGHT

© 2022 Tirnaz, Zandberg, Thomas,
Marsh, Edwards and Batley. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

Application of crop wild relatives in modern breeding: An overview of resources, experimental and computational methodologies

Soodeh Tirnaz, Jaco Zandberg, William J. W. Thomas,
Jacob Marsh, David Edwards and Jacqueline Batley*

School of Biological Sciences, University of Western Australia, Perth, WA, Australia

Global agricultural industries are under pressure to meet the future food demand; however, the existing crop genetic diversity might not be sufficient to meet this expectation. Advances in genome sequencing technologies and availability of reference genomes for over 300 plant species reveals the hidden genetic diversity in crop wild relatives (CWRs), which could have significant impacts in crop improvement. There are many *ex-situ* and *in-situ* resources around the world holding rare and valuable wild species, of which many carry agronomically important traits and it is crucial for users to be aware of their availability. Here we aim to explore the available *ex-/in-situ* resources such as genebanks, botanical gardens, national parks, conservation hotspots and inventories holding CWR accessions. In addition we highlight the advances in availability and use of CWR genomic resources, such as their contribution in pangenome construction and introducing novel genes into crops. We also discuss the potential and challenges of modern breeding experimental approaches (e.g. *de novo* domestication, genome editing and speed breeding) used in CWRs and the use of computational (e.g. machine learning) approaches that could speed up utilization of CWR species in breeding programs towards crop adaptability and yield improvement.

KEYWORDS

pangenome, wild species, modern breeding, *ex situ* resources, *in situ* resources

What can CWRs offer?

The world population is estimated to come close to 10 billion by 2050, while a food gap of more 50% is expected between 2006 and 2050 (Ranganathan et al., 2016). In addition, the growing consequences of climate change, such as increasing weed prevalence and the occurrence of severe disease epidemics and drought

stresses (Raza et al., 2019) will lead towards billions of dollars of crop yield losses worldwide (Gregory et al., 2009; Mittler and Blumwald, 2010). The IPCC (2014) has projected yield losses of up to 25% due to climate change if crop adaptation and improvement are not implemented (IPCC, 2014). At the same time, diets are changing, with shifting nutritional demands toward gluten free, plant-based protein and low GI (glycaemic index) products (Gaikwad et al., 2020). As a result, there is an urgent need for plant breeders to develop new traits in addition to agronomically important traits such as disease resistance, drought tolerance, and yield improvements. On top of these challenges, the effect of the recent COVID-19 pandemic on future agricultural industries has likely added financial strain to both production and distribution chains due to restricted food trade policies and closure of food production facilities (Aday and Aday, 2020). These factors put farmers in a precarious position, with growing pressure to increase production, while they are placed in an increasingly vulnerable position to crop failure and infrastructure setbacks.

Providing breeders access to diverse genetic resources is essential to facilitate, accelerate and optimise crop improvement approaches while domestication bottlenecks have also restricted modern breeding populations (Allaby et al., 2019). The reduction in genetic diversity induced by domestication bottleneck is well documented among many crops such as common bean (Gepts et al., 1986; Papa and Gepts, 2003). Compared to the domesticated population, there are tremendous genetic diversity persists among crop wild relatives (CWRs). The structure of genetic diversity among wild populations appears to be stronger than domesticated; for example in common bean, the diversity of domesticated beans showed limited geographical structure and much less differentiation among populations and regions while in wild bean population even geographically-short-distanced populations carry significant genetic diversity (Papa and Gepts, 2003). As a result, the addition of CWRs to the current breeding programs can significantly widen the source of genetic variation and selection towards yield, resistance and nutritional quality improvement in crops. CWRs can be defined as any taxon belonging to the same genus as a crop; however this definition will include species that are both closely or remotely related to crops (Maxted et al., 2006). In a narrower definition CWRs belong to the same genus of the crop and are closely related to the crops (i.e they are ranked as same the species or same subgenus) (Maxted et al., 2006; Perrino and Perrino, 2020). Advances in breeding techniques, such as genome sequencing, pangenome construction and *de novo* domestication, have been facilitating traits/gene selection from both closely and remotely, related species where fertility and compatibility will be a barrier in traditional breeding approaches, related CWRs to crops. There are a number of successful examples of CWRs application in breeding, such as disease and pest resistance improvement in wheat, rice, potato, tomato, cassava,

sunflower, banana and lettuce; yield improvement in wheat and rice; and improving tolerance to abiotic stress in rice, tomato, barley and chickpea (Hajjar and Hodgkin, 2007). CWRs have also contributed beneficial traits related to ideal plant architecture and weed suppression in rice (Inagaki et al., 2021).

The diversity among CWRs could also be used to decrease the rate of gene/genetic erosion, which has been happening over decades of crop domestication and intense breeding (Schouten et al., 2019). The FAO estimates that ~75% of the genetic diversity in crop varieties has been lost over the past century (FAO, 1999; Khoury et al., 2022). Genetic erosion restricts breeders by limiting sources of selection for identifying desirable agronomic traits. For instance, 96% of peas grown in the US originated from only 9 varieties (Esquinas-Alcázar, 2005). This limited genetic pool will significantly decrease diversity for natural and artificial selection, and intensify the vulnerability of modified varieties to rapid climate changes and new environmental stresses (Esquinas-Alcázar, 2005). Pangenomic analyses in soybean also revealed a reduction in mean gene count per individual due to domestication (Bayer et al., 2022), with disproportionately high levels of biotic and abiotic stress genes lost in modern breeding populations compared to CWRs (Liu et al., 2020). Fortunately, the application of wild species in breeding programs can be used to recover lost diversity caused by erosion, and boost diversity among the crops. SNP array analysis showed that genetic diversity among commercial tomato varieties (from NW Europe) increased by a factor of eight over 7 decades (starting from the 1950s) as a result of the introgression of many disease resistances genes from wild relatives (Schouten et al., 2019).

The application of CWRs in breeding has been also shown to deliver huge economic returns in agricultural industries worldwide, with their annual contribution to the world economy estimated at around US \$186.3 billion in 2020 (Tyack et al., 2020; Bohra et al., 2022). It has been estimated that around 30% of crop yield improvement since 1945, valued worldwide at around US \$100 billion, is a result of CWR use in crop breeding (Pimentel et al., 1997; Brozynska et al., 2016). In tomato, one wild variety provided genes increasing solids content by 2.4% which was worth US\$250 million a year to the global tomato industry; and genes from three wild peanut varieties increased resistance to the root knot nematode, for potential savings of around US \$100 million each year worldwide (Maxted, 2008).

Despite all the potential that CWRs can offer to improve breeding programs, their *in-situ* (in their natural habitats) and *ex-situ* (outside their natural habitats) conservation has been neglected over many years, leading to their potential extinction. Global and local studies have been conducted to guide CWR conservation strategies and estimate the potential loss of diversity of CWRs if the required actions have not been taken. In the US, conservation assessments for 600 CWRs show 42 taxa

(7%) are critically endangered in their natural habitats, 297 (50%) are endangered, 166 (28%) are vulnerable, 66 (11%) are near threatened, and only 23 (3%) are of least concern (Khoury Colin et al., 2020). Another CWR conservation study revealed that the diversity of CWRs is poorly represented in genebanks while out of 1,076 taxa related to 81 crops, for 313 (29%) taxa no germplasm accessions exist, and for 257 (23%) taxa fewer than ten accessions exist (Castañeda-Álvarez et al., 2016). A conservation study on 29 threatened CWRs in Italy, also indicates 23 out of 29 species, have no gene pool at all. In addition, there is not enough data of their *ex-situ* and *in-situ* conservation while 16 and 22 species were identified as high priority for *ex-situ* and *in-situ* conservation respectively (Perrino and Wagensommer, 2022).

Rapid advancements in sequencing technology and computational approaches offer excellent opportunities to fully harness CWR diversity for crop improvement. However, the availability and accessibility of the existing CWR genebank and germplasm resources, capability of modern breeding methodologies and techniques in use of CWRs conservation strategies are currently not well developed to support their full potential and contribution in the current breeding programs. In this regard, here we discuss available *in-/ex-situ* resources for the preservation of CWR variation and the advances in the modern experimental methodologies and computational tools to facilitate capturing the genetic diversity among CWR and their utilization in breeding.

Ex-situ resources

Ex-situ resources, e.g. genebanks and botanical gardens, facilitate user access to plant samples without the need for collecting samples directly from their natural habitat, which can be laborious and complicated when species only exist in remote locations and in most cases need collecting permit (PolicyReport, 2016) and in many cases may not be accessible because of political or socio-economic unrest. The number of accessions held worldwide in genebanks estimated at ~7.4 million accessions in 2009, which increased more than 1.4 million from 1996, ~30% of this increase associated with CWR (van Bemmelen van der Plaats et al., 2021). There are now more than 1750 genebanks worldwide, with 130 of them holding more than 10,000 accessions each (Bohra et al., 2021). Wheat (856,168 accessions), rice (773,948 accessions), barley (466,531 accessions), maize (327,932 accessions) and bean (261,963 accessions) are the most represented crops across the world's genebanks (Wambugu et al., 2018).

To facilitate global access and the conservation of genetic diversity of cultivated and CWR species, genebanks work collaboratively; for instance, Genesys is a database (platform) that contains information of around 4 million accessions across 450 institutes and allows researchers, breeders and policymakers to

browse across all genebanks (<https://www.genesys-pgr.org/content/about/about>) (Table 1). The Genesys database also includes accession information of three of the world's largest genebank databases; the Consultative Group on International Agricultural Research (CGIAR), European Search Catalogue for Plant Genetic Resources (EURISCO), and the U.S. National Plant Germplasm System (NPGS). In contrast to CGIAR and EURISCO that hold both crops and CWRs accessions, the NPGS collection mainly focuses on crop germplasm (<https://www.ars-grin.gov/Pages/Collections#bkmk-1>). The EURISCO database contains over 2 million accessions of crop plants and their wild relatives preserved *ex situ* by about 400 institutes (https://eurisco.ipk-gatersleben.de/apex/eurisco_ws/r/eurisco/home). CGIAR is a partnership of 11 genebanks conserving over 700,000 accessions of cereals, grain legumes, forages, tree species, root and tuber crops and banana and their wild relatives (Table 1). For instance, one of the CGIAR genebank partners is the International Institute of Tropical Agriculture (IITA) which holds over 28,000 accessions of plant material or germplasm of major African crops, including cassava, plantain and banana, yam, soybean, bambara ground-nut and maize. IITA holds the world's largest collection of cowpeas, with 15,122 samples from 88 countries, representing almost half of the global diversity (<https://www.iita.org/research/genetic-resources/>). There are also several genebanks that hold local genetic diversity of crop wild relatives, for example, the Karlsruher Institute of Technology (KIT) collected around 250 species of CWRs with 4500 accessions from all over Germany (<https://www.botanik.kit.edu/garten/english/1056.php>) (Table 1).

Recourses available in genebanks have been used in a number of studies, for example Abdallah et al. (2020) obtained 285 accessions, representing 13 *Lathyrus* (grass pea) species, from The International Center for Agricultural Research in the Dry Areas (ICARDA) and showed that wild *Lathyrus* species have higher resistance to broomrape weeds (*Orobanche* spp.), a root holoparasitic plant that causes significant damage to legume crops (Abdallah et al., 2021). Dida et al. (2021) obtained 52 finger millet accessions, including landraces, wild lines and hybrids between wild and cultivated genotypes, from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) and Genetic Resources Research Institute (GeRRI) genebanks and found that wild accessions were more resistant to blast disease, caused by the *Magnaporthe grisea* fungus, in comparison to the cultivated accessions (Dida et al., 2021).

In addition to the germplasm conservation, there are also genebanks that provide seed kits to smallholder farmers to improve local access to the crop diversity towards better nutrition and supporting climate-resilient agriculture these also assist with the improvement of local genetic diversity among crops. For example, the World Vegetable Center (WorldVeg) genebank distributed over 42,000 seed kits, containing over 183,000 vegetable seeds, to smallholder farmers in Tanzania, Kenya and Uganda, between 2013 and 2017. The kits contained seed of promising accessions and open-

TABLE 1 A list of well-known resources and platforms that include CWR species.

Type	Platform/ resource	Details	Reference/Access
Ex situ: Genebanks	KIT	4500 accessions of CWR across Germany	https://www.botanik.kit.edu/garten/english/1056.ph
	APG	70,000 accessions of pasture and forage	https://pir.sa.gov.au/research/australian_pastures_genebank
	World Vegetable Center	Holding 64,948 accessions which represent 330 species from 155 countries. Species are globally important vegetables such as tomato, onion, peppers, and cabbage as well as more than 10,000 accessions of traditional vegetables	https://avrdc.org/
	Genesys	Partnership with the 3 following genebanks	https://www.genesys-pgr.org/
	&Eurisco	Includes hundreds of research centers, genebanks and institutions across Europe	https://eurisco.ipk-gatersleben.de/apex/?p=103:1
	&NPGS	Includes information of genebanks across U.S., managed by United States Department of Agriculture (USDA). NPGS comprises 20 separate institutions involved in plant germplasm collection, preservation, and distribution.	https://www.ars-grin.gov/Pages/Collections
	&CGIAR	Partnership with the following 11 genebanks	
	&African Rice Center	Holding almost 22,000 rice accessions, 85% of which originated in Africa	https://www.genebanks.org/genebanks/africanrice/
	&Bioversity International	Holding the world's largest collection of banana diversity, including more than 1,500 accessions of cultivated and wild species	https://www.genebanks.org/genebanks/biodiversity-international/
	&CIAT	Holding diverse collections of beans and tropical forages as seed and whole plants, and cassava <i>in vitro</i> and as small plants	https://www.genebanks.org/genebanks/ciat/
	&CIMMYT	Holding large collections of maize (28,000 accessions), including wild teosinte and <i>Tripsacum</i> wild relatives of maize and large collection of wheat (150,000 accessions) including landraces and wild relatives.	https://www.genebanks.org/genebanks/cimmyt/
	&CIP	The world's largest collection of Potato and sweet potato and contains nearly all of the potato wild relatives	https://www.genebanks.org/genebanks/international-potato-centre/
	&ICARDA	Holding diverse collections of barley and wheat, grain legumes and forages, mostly traditional landraces and wild species from the Fertile Crescent	https://www.genebanks.org/genebanks/icarda/
	&ICRAF	Holding 190 species of wild, partially domesticated and domesticated trees	https://www.genebanks.org/genebanks/icraf/
	&ICRISAT	Holds more than 123,000 accessions of cultivated and wild relatives of pulses and cereals, including chickpea, sorghum and pigeonpea	https://www.genebanks.org/genebanks/icrisat/
	&ILRI	Collection of nearly 20,000 forage accessions, of which 97% are wild species	https://www.genebanks.org/genebanks/ilri/
	&IRRI	The largest collection of rice diversity in the world, with more than 130,000 accessions, including genetic stocks, landraces and wild relatives	https://www.genebanks.org/genebanks/irri/
&IITA	Holding a collection of important African crops, including Bambara groundnut, cowpea, maize, soybean	https://www.genebanks.org/genebanks/iita/	
Ex situ: Botanical gardens	BGCI	Including various databases such as PlantSearch, GardenSearch, ThreatSearch and GlobalTreeSearch	https://tools.bgci.org/garden_search.php
	BGCI-US	Botanic Gardens Conservation International and United States Botanic Garden collaboration effort identifying 22 global and 108 priority CWRs not reported in crop gene banks.	(Meyer and Barton., 2019)
<i>In situ</i>	ECPGR	Comprehensive concept for <i>in situ</i> conservation of CWR in Europe	https://www.ecpgr.cgiar.org/
	National Parks in India	Complete list of 103 national parks found in India currently under the protection of the Government	https://www.careerpower.in/national-parks-india.html
	RBGSYD	Australian <i>in situ</i> conservation site containing several key CWRs such as Macadamia nut, finger lime, etc	https://www.rbgsyd.nsw.gov.au/
	UNESCO	Biosphere reserves for the promotion of conserving biodiversity with sustainable use	https://en.unesco.org/biosphere
Software/ tools	GRIN-Global	Genebank information management system (open-sourced software)	(Postman et al., 2009) https://www.grin-global.org/
	CWPs	CWR phylogenetic classification system	(Viruel et al., 2021)
	plaBiPD	Phylogenetic relations among flowering plants with published genome sequences	https://www.plabipd.de/plant_genomes_pa.ep

(Continued)

TABLE 1 Continued

Type	Platform/ resource	Details	Reference/Access
	Mercator	plaBiPD associated protein annotation tool	https://www.plabipd.de/mercator_about.html
	Nordic	Provides towards the planning and implementation of active <i>in-situ</i> and <i>ex-situ</i> conservation of CWR at a national level	http://www.cropwildrelatives.org/conservation-toolkit/introduction/
	GBIF	The Global Biodiversity Information Facility is a global database for the distribution of crop wild relatives with over 5 million records (34% germplasm records, 66% herbarium records)	https://www.gbif.org/dataset/07044577-bd82-4089-9f3a-f4a9d2170b2e
Genomic databases	ECP-GR Natura 2000	A tool for protected area managers of Europe to help identify which CWR genera are likely to occur in the protected areas	https://www.ecpgr.cgiar.org/crop-wild-relatives-in-natura-2000
	PLAZA – Monocots	Access to whole genome sequencing data	https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v5_monocots/
	PLAZA - Dicots	Access to whole genome sequencing data	https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v5_dicots/
	Germinate	A generic plant genetic database with several CWRs	https://germinateplatform.github.io/get-germinate/
	NCBI	General database for literature, genes, genomes and protein sequences of CWRs (Not-CWR specific)	https://www.ncbi.nlm.nih.gov/
	CerealsDB	SNP database for wheat	https://www.cerealsdb.uk.net/cerealgenomics/CerealsDB/indexNEW.php
	Brassica Information Portal (BIP)	Brassica specific	https://www.brassica.info/
	Genome Database for Rosaceae (GDR)	Rosaceae specific	https://www.rosaceae.org/
	GenRes Gateway	Access point to the European genetic resources for plants, forests and animals (including CWR)	https://www.genres.eu/
	ECPGR Central crop database	A database containing passport data, characteristics and primary evaluation data of the major collections of the respective crops.	https://www.ecpgr.cgiar.org/resources/germplasm-databases/ecpgr-central-crop-databases

International Center for Tropical Agriculture (CIAT), International Maize and Wheat Improvement Center (CIMMYT), International Potato Center (CIP), International Center for Agricultural Research in the Dry Areas (ICARDA), World Agroforestry (ICRAF), International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), International Livestock Research Institute (ILRI), International Rice Research Institute (IRRI) and International Institute of Tropical Agriculture (IITA), European Search Catalogue for Plant Genetic Resources (EURISCO) and the National Plant Germplasm System (NPGS), Botanic Gardens Conservation International (BGCI), Crop wild phylorelatives (CWPs), The International Center for Agricultural Research in the Dry Areas (ICARDA), NCBI – National Center for Biotechnology information.

pollinated breeding lines of traditional African vegetables, tomato, Capsicum pepper and soybean. The results show that introduced diversity through seed kits effectively improved local nutrition by facilitating access to various vegetables and also the introduction of new germplasm may slow down genetic erosion and enhance local vegetable diversity (Stoilova et al., 2019).

One of the main concerns across genebanks is the misclassification of species, as previously species identification was mostly based on morphological traits. However, recently the combination of traditional methods combined with molecular approaches, such as DNA barcoding, have improved the accuracy of species identification (van Bemmelen van der Plaats et al., 2021). For example, Mason et al., 2015 proved high-throughput genotyping approaches, such as a SNP array, is an effective methodology for species confirmation. They performed diversity assessment, using the Illumina Brassica 60K SNP array, across 180 Brassicaceae samples sourced from the Australian Grains Genebank and showed 76 of samples were misclassified (Mason

et al., 2015). Through advances in genome sequencing technology and introduction of marker assisted breeding, the use of CWRs has intensified and with this growing interest it is important to keep information in the genebanks well documented and accurate. This is particularly important for use of CWRs in breeding programs where the success rate is highly dependent on the genetic distance between the species, particularly in approaches where crossing compatibility is important, it is crucial to have accurate information regarding the species taxonomy.

Botanical gardens are another *ex-situ* resource for germplasm; moreover they play a crucial role in preventing species extinction through integrated conservation actions (Mounce et al., 2017). Mounce et al. (2017) showed that botanic gardens contribute to the conservation of at least 105,634 species, representing 30% of all plant species diversity, including over 41% of known threatened species (Mounce et al., 2017). The Botanic Gardens Conservation International (BGCI) has the largest collection of living plants (Table 1). The

GardenSearch database, within BGCI, is the only global source for botanical gardens and includes information on over 3,755 botanical institutions worldwide. GardenSearch allows users to search botanical gardens based on their location (country) and their specific features or expertise (https://tools.bgci.org/garden_search.php). For example, based on information stored in GardenSearch, the botanical garden of South Australia has a collection of 40% of Australian flora including drought and salt tolerant plants. This information can facilitate the access and identification of plants with traits of interest for both breeding and research purposes. PlantSearch within the BGCI searches across 1,582,767 collection records, representing 642,718 taxa, at 1,194 institutions; in addition with Plant Search there is a specific option for CWR search at the taxa level (https://tools.bgci.org/plant_search.php) (Mounce et al., 2017).

In-situ resources

In contrast to *ex-situ* conservation sites, *in-situ* sites are typically natural habitats which are rarely curated, for example conservation/rehabilitation facilities or national parks. The benefit of *in-situ* resources is that they are genetically dynamic and continue to evolve in response to both natural and artificial selection, thereby enhancing their adaptation to the environments in which they are grown (Phillips et al., 2016). However, these *in-situ* collections are vulnerable to habitat destruction and/or encroachment caused by civil strife, human settlement pressure and natural disasters including wildfires, flooding, drought and volcanic eruptions. As such, the development of effective CWR conservation strategies is required nationally and globally. Several nations have already prioritised *in situ* CWR conservation, for example, Cyprus (178 priority CWR taxa) (Phillips et al., 2014), UK (148 priority CWR taxa) (Maxted et al., 2007; Fielder et al., 2015), US (821 priority CWR taxa) (Khoury et al., 2013; Khoury et al., 2019), Mexico (310 priority CWR taxa) (Contreras-Toledo et al., 2018), Czech (238 priority CWR taxa) (Taylor et al., 2013) and Norway (204 priority CWR taxa) (Phillips et al., 2016). These *in-situ* conservation efforts provide an ongoing roadmap for the study of the evolutionary history of the plant, which can provide insight into the persistence of traits, identification of new agriculturally significant traits and maintaining biodiversity (Khoury et al., 2020). However, the incorporation of CWRs into traditional farming systems must be carefully considered as it may lead to unfavourable outcomes, for example, a study by Bernal et al., 2019., found that by incorporating a secluded maize genotype (*Zea diploperennis*) into Mexican and Argentinian farms, the pest ‘corn leafhopper’ was able to emerge as a widespread pest to corn farmers (Bernal et al., 2019).

Furthermore, CWR *in-situ* sites typically overlap with regions of high biodiversity, for example, as described by Vincent et al. (2022), the identified Mediterranean basin CWR hotspot shared 91% of its area with a region of high biodiversity, similarly, the California Floristic Province shared 90% between

the CWR and biodiversity hotspots. This overlap has since been harnessed to aid in crop diversity and improvement studies, for example, the Unesco biosphere reserves promote solutions that reconcile the conservation of biodiversity with sustainable development (Benz et al., 2000). However, it is important to consider that *in-situ* resources should not only be limited to ‘wild’ regions. Traditional farming systems are not closed and isolated from gene flow, Louette et al., 1997., showed that the maize varieties cultivated by farmers of Cuzalapa, Mexico, changes in composition over time (Iltis et al., 1979; Louette et al., 1997). Despite certain changes to the germplasm being permanent, for example, the teosinte germplasm in maize which persists during advanced generations of backcrossing (Kato and Sanchez, 2002). In addition to the biodiversity hotspots, centers of origin/diversity, defined as global crop domestication regions including high diversity of both crops and their wild relatives (Vavilov, 1926), could be used as major sources for identification of CWRs. These diversity centers/regions include China; India; Indo-Malayan; Inner Asiatic; Mediterranean; Ethiopian; Central American; the Peruvian-Ecuadorian-Bolivian center, with sub-centers in both Chiloe, Chile and around the Brazil-Paraguay border (Vavilov et al., 1992; Pironon et al., 2020; Maxted and Vincent, 2021). Recently, by assessing the distribution of 222 major international crops and 2,731 of their wild relatives, including both closely and distant related wild species to the crops, Pironon et al. showed geographic distribution of major crop species and their closely related wild species strongly overlap with the Vavilov centers (Pironon et al., 2020). Identification of both crop and wild species diversity hotspots will provide opportunities for identifying and applying more focused conservation strategies for CWRs.

Considering CWRs have been neglected for years and there are many endangered species assessment of national and/or global *in-situ* resources to identify which CWRs are endangered or becoming extinct, whilst screening areas that are rich in wild crops and biodiversity (Hübner and Kantar, 2021) is crucial for protecting CWRs. For example, an assessment of wild banana species (*Musa* spp.) found that 11 out of 59 CWRs are vulnerable and another nine are endangered (Mertens et al., 2021). Khoury et al. (2019), found that of 600 CWR taxa assessed 7% may be critically endangered in their natural habitat and 50% may be endangered. These assessment programs involve a ‘gap analysis’ whereby the currently known and available CWR taxa (*in-situ/ex-situ* resources) are evaluated for their ability to provide future biodiversity to improve food security (Zair et al., 2021). By conducting a thorough gap analysis, Ng’uni et al., 2019., found that 459 CWR taxa out of a national Zambian inventory of 6305 taxa should now be included as part of their conservation and sustainability CWR checklist, with 59 to be specifically prioritised for future food security. The identified taxa represented an agriculturally significant group that was selected due to a shift in socio-economic values to ensure the nation’s food security in the oncoming years. Several nations have conducted their own gap

analysis to ensure food security (Contreras-Toledo et al., 2019; Ng'uni et al., 2019; Tas et al., 2019; González-Orozco et al., 2021; Khaki Mponya et al., 2021; Rahman et al., 2021) and globally ten new *in-situ* conservation sites have been recommended as conservation zones to help achieve global food demand by expanding the *in-situ/ex-situ* resources (Zair et al., 2021).

To successfully establish *in-situ/ex-situ* resources to maintain and improve biodiversity, nations must create an inventory of all known plant taxa. These inventories provide a preliminary resource for the identification of critical taxa, such as CWRs (Teso et al., 2018; Allen et al., 2019; El Mokni et al., 2022). Whilst it is important for each nation to conduct an internal inventory, an unbiased global-scale inventory is also critical to establish CWR taxa. Vincent et al. (2013), originally created a global inventory of important CWR taxa, totaling 1667 taxa, divided between 37 families and 108 genera (Vincent et al., 2013). These inventories serve as the foundation for *in-situ/ex-situ* conservation, as they represent a 'living' CWR databank. However, as these taxa are truly wild, they will continue to evolve, and as such inventories only represent a snapshot of the population from the time of sampling, and recurring sampling is required to update inventories. A list of major global and national inventories is shown in Table 2.

Platforms: Tools for accessing, managing or utilising CWR data and metadata

Several platforms have begun to emerge with the explicit purpose of user-friendliness, designed to aid breeders and scientists alike (Raubach et al., 2021) to facilitate accessibility to CWR resources, including germplasm and genomic data (Table 1). These platforms attempt to solve the most common challenges in handling high throughput data from phenotyping to genotyping: 1) data format, 2) data sharing, 3) data versioning, and 4) historical

data (Raubach et al., 2021). For example, GRIN-global (<https://www.grin-global.org/>) is open-source software for genebank workers to create and manage a genebank's data. Genesys and CGIAR are also examples of genebank platforms/databases (as discussed in the *ex-situ* section) that have been developed at a global scale to efficiently store and categorise data and facilitate the access and conservation of plant species including CWRs. Several other platforms are also available (discussed in the following sections) for visualizing, managing, accessing and storing large datasets related to crops and their relatives.

Software/tool-based platforms

Software/tool-based platforms are essential for data visualisation or organisation and help to gain a better understanding of the accessions stored in genebanks. For example, the Crop wild phylorelative platform (CWP in Table 1) (Viruel et al., 2021) helps to predict the phylogenetic distance (through housekeeping genes or whole genome analysis) and cytogenetic compatibility for breeding programs to help estimate the CWR gene pool classification (Brozynska et al., 2016; Viruel et al., 2021). Alternatively, plaBiPD provides an online platform that visualizes the phylogenetic relationship of genome sequences of flowering plants including CWRs. Furthermore, the associated Mercator online tool allows for the assignment of functional annotations to land plant protein sequences (Schwacke et al., 2019; Bolger et al., 2021).

Database management platforms

Database management tools provide a quick and easy to use platform for the access, management and use of data derived from breeding programs, research studies and trait identification programs using both CWRs and farmed crops. The genotyping platform Germinate v3 (Table 1) (Shaw et al., 2017; Raubach

TABLE 2 A list of major global and national CWR plant inventories.

Inventory name and details	Location assessed	Reference
Globally important CWR taxa	Global	(Vincent et al., 2013)
Important CWR taxa of Mexico	Mexico	(Contreras-Toledo et al., 2018)
National inventory of CWR in Spain	Spain	(Teso et al., 2018)
National inventories of CWR	Portugal	(Brehm et al., 2008)
CWR in USA	USA	(Khoury et al., 2013)
Enhancing and stating the UK CWR inventory	UK	(Fielder et al., 2015)
Prioritised CWR inventory of Italy	Italy	(Landucci et al., 2014)
Enhancing the CWR inventory of Scotland	Scotland	(Fielder et al., 2016)
Setting conservation priorities for CWR in the Fertile Crescent	Fertile Crescent	(Zair et al., 2018)
Prioritised inventory for Tunisia	Tunisia	(El Mokni et al., 2022)
CWR inventory of South, West and North Africa	South, West and North Africa	(Lala et al., 2018; Allen et al., 2019; Nduche et al., 2021)

et al., 2021) provides a rapid directory for importing and exporting plant genetic data such as germ plasm, markers, traits and locations. Germinate v3 has showcased its usefulness in breeding efforts that involve CWRs, specifically those associated with the Crop Trust Crop Wild Relatives project (<https://www.cwrdiversity.org>). Currently, Germinate v3 (20th of April, 2022) contains the directories for CWR taxa: Cowpea (~13100 germplasms), Finger Millet (~1600 germplasms), Grass Pea (~5600 germplasms), Pigeonpea (~2900 germplasms), Chickpea (~23500 germplasm), Alfalfa (~2700 germplasms), Carrot (248 germplasms), Pearl Millet (~2400 germplasms), Barley (~33200 germplasms), Wheat, Sorghum (~2800 germplasms), Eggplant (~3300 germplasms), Rice (~4900 germplasms) and Sunflower (~7900 germplasms) and DIIVA (~2900 germplasms). The use of Germinate has been employed in recent CWR studies. For example, Kouassi et al., 2021., generated interspecies hybrids with eggplants and nine related CWRs. The successfully generated hybrid lines were genotypically and phenotypically screened, wherein it was established that the drought tolerance traits were controlled by genes that are in linkage disequilibrium or have pleiotropic effects. The phenotypic characteristics have been stored in Germinate to provide access to both the user and breeders (Kouassi et al., 2021). Furthermore, Germinate also provides evaluation data of breeding programs. Metwally et al., 2021., generated 13 new superior F₁₀ lines of cowpea by crossing CWRs, improving seed yield and seed quality, as well as introducing earlier maturation. The two datasets which cover 11 different traits for 15 cowpea accessions (total of 2640 data points) were uploaded to Germinate for visualization or downloads (Metwally et al., 2021).

Breeding and research resources are widely available for several crop species such as GrainGenes for wheat, barley, rye and oat (Blake et al., 2019), MaizeGDB for maize (Portwood et al., 2019) and SoyBase for soybean (Grant et al., 2010). These databases primarily host and facilitate the exploration of detailed breeding, pedigree, QTL and molecular information across crop populations. Whilst genomic information regarding CWRs may be presented in these databases, particularly in the case of family-wide databases such as the Sol Genomics Network for Solanaceae (Fernandez-Pozo et al., 2015), they are deposited with no tools for comparative analysis. The development of integrated tools accessible in comprehensive databases is needed to facilitate direct comparisons between wild and domesticated individuals.

Genomic databases

The PLAZA platform holds genomic data of both monocots and dicots. This platform compares the genomic data of submitted dicots and monocots to centralized genomic databases (Van Bel et al., 2022). The submitted genomic data is represented as an interactive phylogenetic tree style figure that links to a bioinformatic

‘workbench’. The workbench includes tools such as gene family plots, collinearity statistic tools, localization tools and direct BLAST tools to the PLAZA protein sequences. Similarly to PLAZA, CerealsDB is a specific database platform for cereals like wheat (Wilkinson et al., 2020), providing several key features such as a SNP database for Axiom[®] 820K and 35K SNP arrays, KASP probes, iSelect Arrays, TaqMan[®] probes. The database is curated to provide agronomically important SNPs (e.g. flowering time associated markers). Furthermore, database platforms such as the Brassica information portal (Brassicaceae) (Eckes et al., 2017) and the Genome database for Rosaceae (Rosaceae) (Evans et al., 2013) have been established as a way to collate and exchange open source information relating to the Brassica and Rosaceae genomes and genetics, respectively, although the databases do not contain CWR resources directly, many of the projects included do include CWR resources. The Legume Information System and Legume Federation project provides an excellent collection of genomic and variant data for over 15 crop species, with a large range of accompanying CWR data (Dash et al., 2016).

Platform models that assist in data handling

A major issue in integrating informatics is a standardised model for data handling, especially as the information regarding the CWR conservation status and breeding programs is diverse and dispersed (Moore et al., 2008). These challenges can be identified by understanding the findable, accessible, interoperable and reusable (FAIR) curation and annotation of minor and underutilized crops (Andrés-Hernández et al., 2021). To address this, the European Crop Wild Diversity Assessment and Conservation Forum developed the Crop Wild Relative Information system (CWRIS) that incorporates an eXtensible Markup Language schema to aid data sharing and exchange. This system integrates with more partitions data into taxon-, site-, and population-specific elements, allowing for the integration with standard conservation biology (Kell et al., 2007; Kell et al., 2008; Moore et al., 2008). CWRIS was developed to provide access of the CWR data to a broader user community such as plant breeders, conservation and rehabilitation site managers, government, biologists and the wider public (Kell et al., 2007). CWRIS has since been integrated into GRIN-Global (<https://npgsweb.ars-grin.gov/gringlobal/taxon/taxonomysearchcwr>), as the website is no longer being maintained or updated.

Pangenomes to capture CWRs genetic variation

In recent years, advances in genome sequencing and bioinformatic tool development have extended the means to fully catalogue genetic variation among domestication and CWR

populations through the construction of pangenomes (Bayer et al., 2020; Jayakodi et al., 2021; Tay Fernandez et al., 2022). Pangenomes achieve this by providing a comprehensive genomic reference to which both small variants, including single-nucleotide polymorphisms (SNPs), and structural variants, including presence/absence variation of large nucleotide sections (PAVs), can be identified across diverse populations (Danilevicz et al., 2020). In addition, analysis of pangenomics allows for the more accurate predication of underlying genetics that are associated with phenotypic variation, such as transposable elements, recombination and double-stranded break/repair (Saxena et al., 2014; Dolatabadian et al., 2020; Song et al., 2020). As pangenomes excel in capturing large structural variation, as is increasingly found between highly divergent populations, they are ideally suited for the comparison of domesticated genomes to CWR taxa to capture 'wild genes' that would be overlooked when using a traditional reference genome (Khan et al., 2020). For example, a pangenome assembly of *Brassica oleracea* with 87 domesticated accessions (Bayer et al., 2021b) identified 58,347 genes across all individuals in comparison to a study that included 8 domesticated accessions and 1 CWR (Golicz et al., 2016) (8 landraces and 1 CWR), which identified a higher number of genes (63,865) (Golicz et al., 2016; Bayer et al., 2021b). Similar findings have been shown in sorghum (Tao et al., 2021) and rice (Xu et al., 2012), where the inclusion of CWR individuals led to large increases in the breadth of genes uncovered.

Beyond capturing more genes, the addition of CWR to pangenomes facilitates the identification of novel SNPs and PAVs that are not found in domesticated populations. For example, Mace et al., 2021 performed comparative analysis in sorghum to quantify the 'contribution of CWR diversity' by establishing the average total number of SNPs per genotype. They found that wild/weedy species contained about one SNP every 763 bp compared to landraces that contained one SNP every 1,282 bp and inbred lines containing one SNP every 1,543 bp (Mace et al., 2021). Lam et al., 2010 also performed a comparative study between 17 wild and 14 cultivated soybean genomes showed higher diversity of SNPs and PAVs among wild species in compared to cultivated. In total, they found 6,318,109 SNPs and 186,177 PAVs, with the CWR genomes carrying 34.66% more SNPs (Lam et al., 2010). This is a clear indication that through optimising our agriculturally important crops, their respective genetic diversity has been reduced and CWR make promises to widen selection diversity (Nelson et al., 2018; Bailey-Serres et al., 2019).

Machine learning and CWRs

The application of machine learning (ML) has proven its efficiency in handling huge amounts of data and is becoming more popular in various plant science fields including gene identification and classification, and biodiversity analysis (Bayer et al., 2021a). For example, in *Arabidopsis* a ML model

was developed to identify candidate stress-related genes by comparing whole genome expression data between the control and stress samples (Wegrzyn et al., 2014). In soybean, a ML model was developed to predict agronomically important traits, including yield, protein, oil, moisture and height, using SNP markers (Liu et al., 2019). Similarly, Ma et al., 2018., successfully developed a ML model to predict eight phenotypic traits among 2000 wheat individuals using 33,709 DArT (Diversity Array Technology) markers (Ma et al., 2018). ML is now also being used to predict mature yield in early development using a combination of image and genotype data (Danilevicz et al., 2021; Danilevicz et al., 2022). Recently ML models were developed for identification of core and dispensable genes in *Oryza sativa* L. and *Brachypodium distachyon* (L.) P. Beauv. using existing pangenomic information. The significant potential of these models is to identify core and dispensable genes in a new species without construction of pangenome (Yocca and Edger, 2022), such approaches can facilitate and speed up genes identification in new cultivated and wild species.

Understanding and usage of environmental conditions, in particular of CWR populations helps in selecting individual populations for the specific introgression goal. CWRs and landraces have occupied local niches (e.g., hot vs. cold regions) and have been shaped by natural selection (Cortés and López-Hernández, 2021), and these traits can be easily tracked when considering collection environmental site parameters. For example, Ariani et al., 2018, by using ~20,000 SNPs across 249 accession of wild *Phaseolus vulgaris*, identified 5 geographically distinct subpopulation, which mostly affected by temperature and rainfall of the regions (Ariani et al., 2018) Berny Mier Y. Teran et al., 2020, also documented that the lines driven from wild parents from the lower rainfall regions produced higher yield in both drought and watered conditions in compare to lines driven from domesticated parents (Berny Mier Y. Teran et al., 2020). Using ML algorithms is also a powerful approach to combine information of germplasm resources and environmental conditions for identification of candidate germplasms with traits of interest. This approach, finding adaptative traits based on environmental parameters, is known as FIGS (Focused Identification of Germplasm Strategy) (Khazaei et al., 2013). Several ML models based on the FIGS approach have been successfully developed and used for identifying germplasm of interest (Table 3). For instance, the identification and classification of *Vicia faba* genetic resources with traits related to drought tolerance (Khazaei et al., 2013). Similarly, in wheat, ML algorithms used for analysing accumulative stem rust trait data (1988-1994), and geographical data of accessions (including landraces and improved accessions) screened for stem rust over 2,000 collection sites revealed an association between the geographic distribution of resistance accessions and environmental variables at collection sites (Bari et al., 2012). Another ML model was successfully developed to predict stripe rust resistance in wheat, based on the stripe rust scores of 725 wheat landrace accessions

with collection site information associated with 2,910 accessions in the ICARDA genebank (Bari et al., 2014). Genetic diversity analysis among 80,000 wheat accessions (including 3,903 wild relatives) also revealed landraces with unexplored diversity and genetic footprints defined by regions under selection (Sansaloni et al., 2020). ML has facilitated the study and discovery of several genetic resources with agronomically valuable traits in crops. There are also “global database for the distribution of wild relatives” (<https://www.gbif.org/dataset/07044577-bd82-4089-9f3a-f4a9d2170b2e>) which includes the distribution data of crop wild relatives that can be used to extract geographical information and potential environmental conditions for CWRs.

Limitation to uses of CWRs within breeding programs

There are many challenges that still prevent the wide-spread use of CWRs as a source of superior alleles that can be incorporated into elite cultivated germplasm. The relatedness,

compatibility and crossability of CWRs to their cultivated counterparts is one issue largely inhibiting the straightforward introduction of CWR traits through traditional breeding. For example, in cotton highly disease resistant sources were identified in wild diploid species, including *Gossypium longicalyx* J.B. Hutch. & B.J.S. Lee; *G. somalense* (Gürke) J.B. Hutch.; *G. stocksii* Mast.; *G. arboreum* L.; and tetraploid species of *G. barbadense* L. (Yik and Birchfield, 1984); however due to genetic incompatibility, ploidy, climbing growth habit, photoperiodism, and agronomic issues breeders were unable to use these resources. Later, through the development of three-species hybrids, researchers were successfully able to introduce donor plants which were fertile and had reniform nematode resistance (Robinson et al., 2004; Konan et al., 2007).

Furthermore, trait identification and selection might be challenging and significantly affected by environment as there are radically different selection regimes in a wild state/region compared to a domesticated state/region while a trait can be useful in a domesticated state (and selected for) may not be useful in the wild and vice-versa. For example, Parker et al.,

TABLE 3 Case studies of the most recent applications of CWRs for crop improvement.

CWR	Application	Outcome	Reference
CWR of Cinnamomum, Piper, Vigna and Oryza in Sri Lanka	ML models to simulate the potential distribution across nine CWR species	The model was able to identify highly vulnerable species to climate change and predict the potential decrease in their suitable habitat by 2050. The study also identifies potential CWR rich areas for future <i>in-situ</i> conservation.	(Ratnayake et al., 2021)
ICARDA genebank barley accessions	FIGS <i>via</i> ML models	Providing predictive characterization for entire ICARDA barely collection	(Azough et al., 2019)
Wild blueberry	ML algorithms for yield prediction by evaluating bee species composition and weather factors	Prediction (with 93% accuracy) showed bee species composition and weather are significant in yield variability while wet rainy springs will greatly reduce blueberry yield.	(Obsie et al., 2020)
Wild cacao	Using ML model for surveying canopy and vegetation assessments	92% of classification accuracy for the structural attributes of the canopy	(Duarte-Carvajalino et al., 2021)
Large collection of <i>Vicia faba</i> L.	ML models used to evaluate FICS approach for identification of traits related to drought	The model was successful to indicate leaflet, canopy temperature and relative water content are important traits for drought-tolerance selection.	(Khazaei et al., 2013)
<i>Solanum pimpinellifolium</i>	Genome editing (<i>de novo</i> domestication)	Produced a modified version of the wild <i>S. pimpinellifolium</i> which displayed a 10 times increase in the number of fruit and a 3 times increase in fruit size. The fruit also contained 500% more lycopene compared to the commonly cultivated <i>S. lycopersicum</i> .	(Ariani et al., 2018)
<i>Physalis pruinose</i>	Genome editing (<i>de novo</i> domestication)	Edited orthologues of cultivated tomato in the distant relative <i>P. pruinose</i> to improve plant architecture, flower production and fruit size.	(Lemmon et al., 2018)
<i>Oryza alta</i>	Genome editing (<i>de novo</i> domestication)	Established the first ever polyploid rice by genome editing the allotetraploid relative <i>O. alta</i> .	(Yu et al., 2021)
<i>Aegilops tauschii</i>	Association genetics with resistance gene enrichment sequencing (AgRenSeq)	Developed the AgRenSeq methodology and identified two novel wheat stem rust resistance genes, <i>Sr46</i> and <i>SrTA1662</i> , in a wild wheat progenitor.	(Arora et al., 2019)
<i>Solanum americanum</i>	Resistance gene enrichment sequencing and single-molecule real-time sequencing (SMRT RenSeq)	Identified the genome-wide repertoire of nucleotide-binding leucine-rich repeat type R genes in the wild <i>S. americanum</i> and cloned <i>Rpi-amr3i</i> , a novel R gene for potato late blight.	(Witek et al., 2016)
<i>Oryza rufipogon</i>	Genome editing	Optimised an efficient transformation system in wild rice, aiding future genome editing efforts including <i>de novo</i> domestication.	(Xiang et al., 2022)
<i>Solanum peruvianum</i>	Genome editing	Developed a genome editing approach using protoplast regeneration for the tetraploid wild tomato relative.	(Lin et al., 2022)

(2020), suggested the decreased-pod dehiscence (PD) trait among domesticated haplotypes of common bean is as a result of the different fitness landscape imposed by domestication, where stronger selection pressure were used against PD in arid condition of North Mexico compared to tropical lowlands (Andes), where environmental humidity masks susceptibility to PD and reducing selection pressure against it (Parker et al., 2020). It is also often challenging to accurately evaluate the yield of CWRs since they can display growth forms or traits that are difficult to manage, for example the wild progenitor of common bean has naturally dehiscent seed pods, making yield measurements arduous to obtain, and has a larger, less compact growth habit that is far less suitable for cultivated environments compared to cultivated common bean (Koinange et al., 1996). Even if beneficial wild derived traits are introgressed into elite material, they can often have a negative effect on yield or yield-related traits, through linkage drag. A common example is the introduction of biotic stress tolerance genes, for example disease resistance genes, which improve some resistance/tolerance but are detrimental to other agronomic traits (Brouwer and St Clair, 2004; Summers and Brown, 2013). Furthermore, after introducing genetic material from CWRs into an elite background, problems with sterility, often seen at the F₁ or BC₁ generation, can arise (Wang et al., 2020; Bohra et al., 2022).

There are also a number of challenges of CWR application in breeding that have been eased by availability of more genomic resources, and advances in laboratory techniques, as discussed in the following section. These include lack of information of gene-trait relationships in wild species, uncertainty of how allelic combinations will be expressed in different cultivated crop backgrounds and difficulties of transferring genes of interest into crops (Dempewolf et al., 2017).

Modern breeding and CWRs

There are now avenues to harness CWRs and overcome some of these barriers. For instance, wild-derived genes conferring desirable alleles can now be introduced through precise genome editing into elite backgrounds without the need for lengthy introgression regimes, bypassing the barriers of linkage drag and reduced fertility that so often complicate the use of CWRs (Bohra et al., 2021). These modern approaches, utilising the advances in genomics and genome editing, provide promising pathways to overcome long-standing challenges and push CWRs to the forefront of crop improvement. Table 3, included examples of successful application of CWRs for crop improvement *via* modern breeding approaches.

Genomics provides an avenue to explore the genetic diversity in CWRs and identify agronomically valuable genes or QTL. Sequencing CWRs followed by *de novo* assembly can generate reference assemblies that underpin downstream

applications, such as the functional characterization of genes and targeted genome editing. Although initially lagging behind cultivated crop genomes, a number of CWRs assemblies are now becoming available, including relatives of barley, rice, soybean, tomato and wheat (Brozynska et al., 2016; Bohra et al., 2022). Often in combination with high-throughput phenotyping, these genome assemblies have enabled the identification of several important genes and QTL from CWRs, for example numerous disease resistance genes in wheat (Yahiaoui et al., 2009; Periyannan et al., 2013; Sainetnac et al., 2013) and QTL associated with oil content in soybean (Zhou et al., 2015). High-quality assemblies based on third generation long read sequencing are now becoming the standard for reference genomes in major crops. Advances in long-read sequencing in terms of increased accessibility and lower price points, will be vital for the construction of high-quality long read assemblies in a broad range of CWRs, which will unlock an arsenal of beneficial CWR genetic diversity ready to be harnessed for crop improvement.

There are also recent genomic methodologies that have been developed to identify genes linked to specific traits; for instance resistance gene enrichment sequencing (RenSeq) is a methodology that targets, enriches and sequences *R* genes within any plant genome based on common *R* gene motifs (Jupe et al., 2013). To date, it has been used to capture nucleotide-binding-site leucine-rich repeat proteins (NLRs), receptor-like proteins (RLPs) and receptor-like kinases (RLKs), which represent the largest families of *R* genes (Jupe et al., 2013; Lin et al., 2020). Since its initial development, RenSeq has been combined with other approaches, including ethyl methanesulfonate (EMS) mutagenesis (MutRenSeq), single-molecule real-time sequencing (SMRT RenSeq) and association genetics (AgRenSeq). These combined workflows have rapidly identified and cloned causative *R* genes in a wild potato relative (Witek et al., 2016), wheat (Steuernagel et al., 2016), wild diploid wheat (Arora et al., 2019) and rye (Vendelbo et al., 2022). RenSeq is a promising alternative to whole genome sequencing for large scale *R* gene identification, and if utilised in CWRs, has the potential to rapidly expand the *R* gene arsenal used for breeding disease resistant cultivars. Notably, AgRenSeq does not rely on a reference genome (Arora et al., 2019), therefore it is extremely applicable to CWRs that are yet to have a reference assembly, but whose cultivated counterpart has well characterised *R* genes.

While there has been rapid progress within the field of plant genome editing, the application within CWRs has been far slower. The limited genomic resources for many CWRs serves as an initial barrier, then the lack of functionally characterized gene targets and easy delivery system for those targets proves arduous. In spite of these challenges, one innovative application of CRISPR recently proposed is the manipulation of genes controlling important agronomic traits, for example plant architecture genes, while purposefully retaining valuable wild-

derived traits such as stress tolerance or improved nutritional quality; in essence, the domestication of a CWR or landrace that has never been cultivated. This approach, termed *de novo* domestication, can produce new crops from a CWR in a matter of generations through genome editing technology (Gasparini et al., 2021). Using a wild tomato relative, Zsögön et al., 2018., edited four key tomato domestication genes, *SELF-PRUNING*, *OVATE*, *FRUIT WEIGHT 2.2* and *LYCOPENE BETACYCLASE*, to produce an engineered tomato crop boasting increased fruit number and size compared to the wild parent, and vastly improved nutritional quality compared to cultivated tomato (Zsögön et al., 2018). A similar approach was undertaken in the orphan crop groundcherry, a distant tomato relative, whereby productivity traits including plant architecture, flower production and fruit size were improved by editing known tomato orthologues with CRISPR-Cas9 (Lemmon et al., 2018). One ambitious study utilised *de novo* domestication to develop the first ever polyploid rice crop, through the rapid domestication of an allotetraploid wild rice, *Oryza alta* (Yu et al., 2021). This has demonstrated a feasible route to create polyploid versions of diploid crops, which are said to benefit from genome buffering *via* gene redundancy, hybrid vigour and environmental fortitude (Mason and Batley, 2015). As researchers characterise more genes related to key domestication traits in model or major crops and high-quality CWR genome assemblies are generated, the potential for editing these genes in CWRs skyrockets, leading to the possible creation of new crops through *de novo* domestication. Furthermore, simultaneously identifying and cataloguing agronomically beneficial traits in CWRs will greatly enhance our ability to exploit wild genetic diversity, meaning *de novo* domesticated crops will be more nutritious and climate resilient than their cultivated relatives.

Despite the promising potential of *de novo* domestication, one of the major challenges preventing the widespread deployment of CRISPR in CWRs, and therefore *de novo* domestication, is the delivery system of the genome editing reagents. Even for elite cultivars, quick and easy methods for delivery that are widely transferable between species remain elusive (Zhan et al., 2021). The most popular DNA delivery approaches include agrobacterium-mediated delivery, which utilises the soil pathogen *Agrobacterium tumefaciens* to transfer DNA into the host genome, and biolistic or micro-projectile-mediated delivery, where the donor DNA is mechanically forced into the host cells (Ran et al., 2017). However, these methods come with certain limitations. *Agrobacterium*-mediated delivery is hindered by its inability to introduce small donor fragments, its difficulty in preventing plasmid integration and thereby producing a transgenic plant, and is dependent on the genotype of the recipient, particularly for monocot plants (Ran et al., 2017). While biolistic methods provide some advantages over *Agrobacterium*-mediated delivery, for example the delivery of multiple targets, its use is lower than expected due to issues with multiple copies of the transgene being incorporated into the host,

resulting in altered gene expression or complete silencing. Efficient delivery methods using these approaches, after significant optimisation, have been established in model plants and select major crops. However, such methods are not easily transferrable to CWRs, as they often represent a diverse set of morphotypes which introduces unique challenges hindering delivery. On top of this, CWRs are also difficult to regenerate, further complicating the transformation process (Zhu et al., 2020).

Several alternative approaches for reagent delivery which were initially developed in animal cells, are being explored in plants (Ghogare et al., 2021). For example, a biolistics approach using nanoparticles offers a less harmful delivery method compared to larger microparticles, which may reduce delivery damage, a common issue encountered in plants due to the presence of a cell wall (Zhang et al., 2019; Cunningham et al., 2020). Most excitingly, delivery mediated by viral vectors can completely bypass the need for regeneration which is an extremely promising prospect for editing hard to regenerate CWRs, however this method is limited by its delivery capacity (Shan-e-Ali Zaidi and Mansoor, 2017). Novel delivery methods will help to overcome the barriers preventing widespread plant transformation and reduce the amount of optimisation needed. In doing so, efficient genome editing in CWRs will be one step closer.

Another potential approach for CWRs utilization in breeding schemes is through speed breeding. The concept of speed breeding revolves around manipulating the photoperiod (e.g. 12 hr extended to 22 hr) and temperature in a controlled growth facility to rapidly produce multiple crop generations per year (Watson et al., 2018). Through speed breeding, the genetic background of cultivars can be fixed in an accelerated timeframe, a process which usually takes years of inbreeding. Speed breeding has been tested and effectively produced multiple generations in a single year for crops such as barley, canola, chickpea, pea, rice, sorghum and wheat (Espósito et al., 2012; Rizal et al., 2014; Watson et al., 2018; Nagatoshi and Fujita, 2019; Rana et al., 2019). In the absence of precise genome editing, desirable traits from CWRs which are introgressed into elite cultivars through traditional breeding will often bring with them unwanted deleterious alleles. Hence, speed breeding can facilitate the quick growth of multiple generations, allowing undesirable traits to be selected against, and for these new varieties to reach a stable genetic background. In addition, speed breeding would benefit alternative approaches to domesticate CWRs without the use of CRISPR, such as germplasm conversion (Stephens et al., 1967; Rosenow et al., 1997; Klein et al., 2016). Germplasm conversion involves the alteration of germplasm through crossing, multiple rounds of selection for various traits and inbreeding to become well-adapted to new environments while also having favourable agronomic traits (Stephens et al., 1967). Extensive germplasm conversion has been done in Sorghum to transform numerous exotic varieties into early-maturing and dwarf-height varieties

that are adapted for cultivation in the US or other temperate regions (Stephens et al., 1967; Rosenow et al., 1997; Klein et al., 2016). As an alternative to genome editing, germplasm conversion could be harnessed to introduce important agronomic traits into CWRs through hybridization and then followed by marker-assisted selection (MAS). The advantage of this over genome editing is that specific knowledge of the target sequences is not required, only knowledge of the genomic region conferring the domestication trait/s. However, it is likely that this method would be more laborious and time consuming compared to genome editing approaches, as several generations are usually required to achieve the final product. Therefore, exposing these CWRs to speed breeding conditions may help to mitigate the time required for cycling multiple generations that is necessary for effective germplasm conversion of CWRs into commercially viable crops (Bhatta et al., 2021).

Conclusion

Crop wild relatives have remained under-utilised during crop domestication and intense crop breeding, despite the fact they harbour beneficial traits such as disease and pest resistance, and tolerance to abiotic stresses. CWRs have the potential to widen selection sources for breeders beyond the existing variation among cultivated crops to meet future foods' quality and quantity demands. A multi-resource integrative approach that utilises many of the resources outlined here will enable CWRs to be effectively used as a source of valuable genetic diversity. For example, ML strategies based on FIGS in combination with genomic and pangenomic resources that capture the gene diversity that exists in CWRs, will help to rapidly identify adaptive traits based on environmental parameters which will in turn guide the identification of genes underpinning these traits. However, realisation and utilisation of the full potential of the genes and diversity presented in CWRs will ultimately depend on the availability of resources and experimental techniques to support breeding programs (Hajjar and Hodgkin, 2007). There are a number of resources and databases that both researchers and breeders can benefit from, but ongoing efforts are crucial to keep these data well organised and up-to-date. This is only possible with the great collaboration between ecological/biological conservation sectors, who manage CWR ex/in-situ conservation and prevent extinction, researchers in the field of computer science, plant biology, for

References

- Abdallah, F., Kumar, S., Amri, A., Mentag, R., Kehel, Z., Mejri, R. K., et al. (2021). Wild lathyrus species as a great source of resistance for introgression into cultivated grass pea (*Lathyrus sativus* L.) against broomrape weeds (*Orobanche crenata* forsk. and *Orobanche foetida* Poir.). *Crop Sci.* 61 (1), 263–276. doi: 10.1002/csc.220399
- Aday, S., and Aday, M. S. (2020). Impact of COVID-19 on the food supply chain. *Food Qual. Saf.* 4 (4), 167–180. doi: 10.1093/fqsafe/fyaa024

example plant genomics and agricultural industries, who assist with identification of traits/genes of interest among CWRs and only with this multidisciplinary effort is there a chance to guarantee the future food demands.

Author contributions

ST and JB conceptualized the review. ST wrote the main text with additions from WT, JZ, JM, DE and JB. DE and JB edited the paper. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the Australian Research Council projects DP200100762, DP210100296 and the Grains Research and Development Corporation (UWA1905-006RTX).

Acknowledgments

WT would like to acknowledge the support of the Grains Research and Development Corporation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Allaby, R. G., Ware, R. L., and Kistler, L. (2019). A re-evaluation of the domestication bottleneck from archaeogenomic evidence. *Evol. Appl.* 12 (1), 29–37. doi: 10.1111/eva.12680

- Allen, E., Gaisberger, H., Brehm, J. M., Maxted, N., Thormann, I., Lupupa, T., et al. (2019). A crop wild relative inventory for southern Africa: A first step in linking conservation and use of valuable wild populations for enhancing food security. *Plant Genet. Resour.* 17 (2), 128–139. doi: 10.1017/S1479262118000515

- Andrés-Hernández, L., Halimi, R. A., Mauleon, R., Mayes, S., Baten, A., and King, G. J. (2021). Challenges for FAIR-compliant description and comparison of crop phenotype data with standardized controlled vocabularies. *Database* 2021, 1–11. doi: 10.1093/database/baab028
- Ariani, A., Berny Mier, Y. T. J. C., and Gepts, P. (2018). Spatial and temporal scales of range expansion in wild phaseolus vulgaris. *Mol. Biol. Evol.* 35 (1), 119–131. doi: 10.1093/molbev/msx273
- Arora, S., Steuernagel, B., Gaurav, K., Chandramohan, S., Long, Y., Matny, O., et al. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nat. Biotechnol.* 37 (2), 139–143. doi: 10.1038/s41587-018-0007-9
- Azough, Z., Kehel, Z., Benomar, A., Bellafkih, M., and Amri, A. (2019). “Predictive characterization of ICARDA genebank barley accessions using FIGS and machine learning,” in *Intelligent environments (Workshops)*, 121–129.
- Bailey-Serres, J., Parker, J. E., Ainsworth, E. A., Oldroyd, G. E. D., and Schroeder, J. I. (2019). Genetic strategies for improving crop yields. *Nature* 575 (7781), 109–118. doi: 10.1038/s41586-019-1679-0
- Bari, A., Amri, A., Street, K., Mackay, M., De Pauw, E., Sanders, R., et al. (2014). Predicting resistance to stripe (yellow) rust (*Puccinia striiformis*) in wheat genetic resources using focused identification of germplasm strategy. *J. Agric. Sci.* 152 (6), 906–916. doi: 10.1017/S0021859613000543
- Bari, A., Street, K., Mackay, M., Endresen, D. T. F., De Pauw, E., and Amri, A. (2012). Focused identification of germplasm strategy (FIGS) detects wheat stem rust resistance linked to environmental variables. *Genet. Resour. Crop Evol.* 59 (7), 1465–1481. doi: 10.1007/s10722-011-9775-5
- Bayer, P. E., Golicz, A. A., Scheben, A., Batley, J., and Edwards, D. (2020). Plant pan-genomes are the new reference. *Nat. Plants* 6 (8), 914–920. doi: 10.1038/s41477-020-0733-0
- Bayer, P. E., Peterleit, J., Danilevicz, M. F., Anderson, R., Batley, J., and Edwards, D. (2021a). The application of pangenomics and machine learning in genomic selection in plants. *Plant Genome* 14 (3), e20112. doi: 10.1002/tpg2.20112
- Bayer, P. E., Scheben, A., Golicz, A. A., Yuan, Y., Faure, S., Lee, H., et al. (2021b). Modelling of gene loss propensity in the pangenomes of three brassica species suggests different mechanisms between polyploids and diploids. *Plant Biotechnol. J.* 19 (12), 2488–2500. doi: 10.1111/pbi.13674
- Bayer, P. E., Valliyodan, B., Hu, H., Marsh, J. I., Yuan, Y., Vuong, T. D., et al. (2022). Sequencing the USDA core soybean collection reveals gene loss during domestication and breeding. *Plant Genome* 15 (1), e20109. doi: 10.1002/tpg2.20109
- Benz, B. F., Cevallos E. J., Santana M. F., Rosales A. J., and Graf, M. S. (2000). Losing knowledge about plant use in the sierra de manantlan biosphere reserve, Mexico. *Economic Bot.* 54 (2), 183–191. doi: 10.1007/BF02907821
- Bernal, J. S., Dávila-Flores, A. M., Medina, R. F., Chen, Y. H., Harrison, K. E., and Berrier, K. A. (2019). Did maize domestication and early spread mediate the population genetics of corn leafhopper? *Insect Sci.* 26 (3), 569–586. doi: 10.1111/1744-7917.12555
- Berny Mier Y. Teran, J. C., Konzen, E. R., Palkovic, A., Tsai, S. M., and Gepts, P. (2020). Exploration of the yield potential of mesoamerican wild common beans from contrasting eco-geographic regions by nested recombinant inbred populations. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00346
- Bhatta, M., Sandro, P., Smith, M. R., Delaney, O., Voss-Fels, K. P., Gutierrez, L., Hickey, L. T., et al. (2021). Need for speed: Manipulating plant growth to accelerate breeding cycles. *Current Opin. in Plant Biol.* 60, 101986. doi: 10.1016/j.pbi.2020.101986
- Blake, V. C., Woodhouse, M. R., Lazo, G. R., Odell, S. G., Wight, C. P., Tinker, N. A., et al. (2019). GrainGenes: centralized small grain resources and digital platform for geneticists and breeders. *Database (Oxford)* 2019, 1–7. doi: 10.1093/database/baz065
- Bohra, A., Kilian, B., Sivasankar, S., Caccamo, M., Mba, C., McCouch, S. R., et al. (2021). Reap the crop wild relatives for breeding future crops. *Trends Biotechnol.*
- Bohra, A., Kilian, B., Sivasankar, S., Caccamo, M., Mba, C., McCouch, S. R., et al. (2022). Reap the crop wild relatives for breeding future crops. *Trends Biotechnol.* 40 (4), 412–431. doi: 10.1016/j.tibtech.2021.08.009
- Bolger, M., Schwacke, R., and Usadel, B. (2021). “MapMan visualization of RNA-seq data using Mercator4 functional annotations,” in *Solanum tuberosum* (New York, NY: Humana), 195–212.
- Brehm, J. M., Maxted, N., Ford-Lloyd, B. V., and Martins-Louçao, M. A. (2008). National inventories of crop wild relatives and wild harvested plants: case-study for Portugal. *Genet. Resour. Crop Evol.* 55 (6), 779–796. doi: 10.1007/s10722-007-9283-9
- Brouwer, D. J., and St Clair, D. A. (2004). Fine mapping of three quantitative trait loci for late blight resistance in tomato using near isogenic lines (NILs) and sub-NILs. *Theor. Appl. Genet.* 108 (4), 628–638. doi: 10.1007/s00122-003-1469-8
- Brozynska, M., Furtado, A., and Henry, R. J. (2016). Genomics of crop wild relatives: expanding the gene pool for crop improvement. *Plant Biotechnol. J.* 14 (4), 1070–1085. doi: 10.1111/pbi.12454
- Castañeda-Álvarez, N. P., Khoury, C. K., Achicanoy, H. A., Bernau, V., Dempewolf, H., Eastwood, R. J., et al. (2016). Global conservation priorities for crop wild relatives. *Nat. Plants* 2 (4), 16022. doi: 10.1038/nplants.2016.22
- Contreras-Toledo, A. R., Cortés-Cruz, M. A., Costich, D., de Lourdes Rico-Arce, M., Brehm, J. M., and Maxted, N. (2018). A crop wild relative inventory for Mexico. *Crop Sci.* 58 (3), 1292–1305. doi: 10.2135/cropsci2017.07.0452
- Contreras-Toledo, A. R., Cortés-Cruz, M., Costich, D. E., de Lourdes Rico-Arce, M., Brehm, J. M., and Maxted, N. (2019). Diversity and conservation priorities of crop wild relatives in Mexico. *Plant Genet. Resour.* 17 (2), 140–150. doi: 10.1017/S1479262118000540
- Cortés, A. J., and López-Hernández, F. (2021). Harnessing crop wild diversity for climate change adaptation. *Genes* 12 (5). doi: 10.3390/genes12050783
- Cunningham, F. J., Demirer, G. S., Goh, N. S., Zhang, H., and Landry, M. P. (2020). “Nanobiologics: An emerging genetic transformation approach,” in *Biological DNA delivery in plants* (New York, NY: Humana), 141–159.
- Danilevicz, M. F., Bayer, P. E., Boussaid, F., Bennamoun, M., and Edwards, D. (2021). Maize yield prediction at an early developmental stage using multispectral images and genotype data for preliminary hybrid selection. *Remote Sens.* 13 (19), 3976. doi: 10.3390/rs13193976
- Danilevicz, M. F., Gill, M., Anderson, R., Batley, J., Bennamoun, M., Bayer, P. E., et al. (2022). Plant genotype to phenotype prediction using machine learning. *Front. Genet.* 13. doi: 10.3389/fgene.2022.822173
- Danilevicz, M. F., Tay Fernandez, C. G., Marsh, J. I., Bayer, P. E., and Edwards, D. (2020). Plant pangenomics: approaches, applications and advancements. *Curr. Opin. Plant Biol.* 54, 18–25. doi: 10.1016/j.pbi.2019.12.005
- Dash, S., Campbell, J. D., Cannon, E. K. S., Cleary, A. M., Huang, W., Kalberer, S. R., et al. (2016). Legume information system (LegumeInfo.org): a key component of a set of federated data resources for the legume family. *Nucleic Acids Res.* 44 (D1), D1181–D1188. doi: 10.1093/nar/gkv1159
- Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C., and Guarino, L. (2017). Past and future use of wild relatives in crop breeding. *Crop Sci.* 57 (3), 1070–1082. doi: 10.2135/cropsci2016.10.0885
- Dida, M. M., Oduori, C. A., Manthi, S. J., Avosa, M. O., Mikwa, E. O., Ojulong, H. F., et al. (2021). Novel sources of resistance to blast disease in finger millet. *Crop Sci.* 61 (1), 250–262. doi: 10.1002/csc.2.20378
- Dolatabadian, A., Bayer, P. E., Tirnaz, S., Hurgobin, B., Edwards, D., and Batley, J. (2020). Characterization of disease resistance genes in the brassica napus pangenome reveals significant structural variation. *Plant Biotechnol. J.* 18 (4), 969–982. doi: 10.1111/pbi.13262
- Duarte-Carvajalino, J. M., Paramo-Alvarez, M., Ramos-Calderón, P. F., and González-Orozco, C. E. (2021). Estimation of canopy attributes of wild cacao trees using digital cover photography and machine learning algorithms. *iForest - Biogeosciences Forestry* 14 (6), 517–521. doi: 10.3832/ifor3936-014
- Eckes, A. H., Gubala, T., Nowakowski, P., Szymczyszyn, T., Wells, R., Irwin, J. A., et al. (2017). Introducing the brassica information portal: Towards integrating genotypic and phenotypic brassica crop data. *F1000Research* 6:465. doi: 10.12688/f1000research.11301.1
- El Mokni, R., Barone, G., Maxted, N., Kell, S., and Domina, G. (2022). A prioritised inventory of crop wild relatives and wild harvested plants of Tunisia. *Genet. Resour. Crop Evol.* 1–34, 1787–1816. doi: 10.1079/9781845930998.0471
- Esposito, M., Almirón, P., Gatti, I., Cravero, V. P., Anido, F. S. L., and Country, E. (2012). A rapid method to increase the number of F1 plants in pea (*Pisum sativum*) breeding programs. *Genet. Mol. Res.* 11 (3), 2729–2732. doi: 10.4238/2012.June.18.1
- Esquinas-Alcázar, J. (2005). Protecting crop genetic diversity for food security: political, ethical and technical challenges. *Nat. Rev. Genet.* 6 (12), 946. doi: 10.1038/nrg1729
- Evans, K., Jung, S., Lee, T., Brucher, L., Cho, I., Peace, C., et al. (2013). Addition of a breeding database in the genome database for rosaceae. *Database* 2013. doi: 10.1093/database/bat078
- FAO (1999) *What is happening to agrobiodiversity?* Available at: <https://www.fao.org/3/y5609e/y5609e02.htm>.
- Fernandez-Pozo, N., Menda, N., Edwards, J. D., Saha, S., Teclé, I. Y., Strickler, S. R., et al. (2015). The sol genomics network (SGN)—from genotype to phenotype to breeding. *Nucleic Acids Res.* 43 (Database issue), D1036–D1041. doi: 10.1093/nar/gku1195
- Fielder, H., Brotherton, P., Hosking, J., Hopkins, J. J., Ford-Lloyd, B., and Maxted, N. (2015). Enhancing the conservation of crop wild relatives in England. *PLoS One* 10 (6), e0130804. doi: 10.1371/journal.pone.0130804
- Fielder, H., Smith, C., Ford-Lloyd, B., and Maxted, N. (2016). Enhancing the conservation of crop wild relatives in Scotland. *J. Nat. Conserv.* 29, 51–61. doi: 10.1016/j.jnc.2015.11.002
- Gaikwad, K. B., Rani, S., Kumar, M., Gupta, V., Babu, P. H., Bainsla, N. K., et al. (2020). Enhancing the nutritional quality of major food crops through

- conventional and genomics-assisted breeding. *Front. Nutr.* 7, 533453. doi: 10.3389/fnut.2020.533453
- Gasparini, K., Moreira, J. D. R., Peres, L. E. P., and Zsögön, A. (2021). *De novo* domestication of wild species to create crops with increased resilience and nutritional value. *Curr. Opin. Plant Biol.* 60, 102006–102006. doi: 10.1016/j.pbi.2021.102006
- Gepts, P., Osborn, T. C., Rashka, K., and Bliss, F. A. (1986). Phaseolin-protein variability in wild forms and landraces of the common Bean (*Phaseolus vulgaris*): Evidence for multiple centers of domestication. *Economic Bot.* 40 (4), 451–468. doi: 10.1007/BF02859659
- Ghogare, R., Ludwig, Y., Bueno, G. M., Slamet-Loedin, I. H., and Dhingra, A. (2021). Genome editing reagent delivery in plants. *Transgenic Res.* 30, 321–335. doi: 10.1007/s11248-021-00239-w
- Golicz, A. A., Bayer, P. E., Barker, G. C., Edger, P. P., Kim, H., Martinez, P. A., et al. (2016). The pangenome of an agronomically important crop plant *Brassica oleracea*. *Nat. Commun.* 7, 13390. doi: 10.1038/ncomms13390
- González-Orozco, C. E., Sosa, C. C., Thornhill, A. H., and Laffan, S. W. (2021). Phylogenetic diversity and conservation of crop wild relatives in Colombia. *Evolutionary Appl.* 14 (11), 2603–2617. doi: 10.1111/eva.13295
- Grant, D., Nelson, R. T., Cannon, S. B., and Shoemaker, R. C. (2010). SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* 38 (suppl_1), D843–D846. doi: 10.1093/nar/gkp798
- Gregory, P. J., Johnson, S. N., Newton, A. C., and Ingram, J. S. (2009). Integrating pests and pathogens into the climate change/food security debate. *J. Exp. Bot.* 60 (10), 2827–2838. doi: 10.1093/jxb/erp080
- Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* 156 (1), 1–13. doi: 10.1007/s10681-007-9363-0
- Hübner, S., and Kantar, M. B. (2021). Tapping diversity from the wild: From sampling to implementation. *Front. Plant Sci.* 12 (38). doi: 10.3389/fpls.2021.626565
- Iltis, H. H., Doebley, J. F., Guzmán M, R., and Pazy, B. (1979). *Zea diploperennis* (Gramineae): A new teosinte from Mexico. *Science* 203 (4376), 186–188. doi: 10.1126/science.203.4376.186
- Inagaki, N., Asami, H., Hirabayashi, H., Uchino, A., Imaizumi, T., and Ishimaru, K. (2021). A rice ancestral genetic resource conferring ideal plant shapes for vegetative growth and weed suppression. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.748531
- IPCC (2014). “Climate change 2014: Synthesis report,” in *Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change*. Eds. R. K. Pachauri and L. A. Meyer (Geneva, Switzerland: IPCC).
- Jayakodi, M., Schreiber, M., Stein, N., and Mascher, M. (2021). Building pangenome infrastructures for crop plants and their use in association genetics. *DNA Res.* 28 (1), dsaa030. doi: 10.1093/dnares/dsaa030
- Jupe, F., Witek, K., Verweij, W., Śliwka, J., Pritchard, L., Etherington, G. J., et al. (2013). Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *Plant J.* 76 (3), 530–544. doi: 10.1111/tj.12307
- Kato, T., and Sanchez, J. (2002). Introgression of chromosome knobs from *zea diploperennis* into maize [*Zea mays* L.]. *Maydica (Italy)* 47(1), 33–5.
- Kell, S., Jury, S., Knüpfner, H., Ford-Lloyd, B., and Maxted, N. (2007). PGR forum: serving the crop wild relative user community. *Bocconea* 21, 413–421.
- Kell, S., Moore, J., Iriondo, J., Scholten, M., Ford-Lloyd, B., and Maxted, N. (2008). CWRIS: an information management system to aid crop wild relative conservation and sustainable use. *Crop wild relative conservation and use* (Wallingford UK: CABI), 471–491. doi: 10.1079/9781845930998.047
- Khaki Mponya, N., Chanyenga, T., Magos Brehm, J., and Maxted, N. (2021). *In situ* and *ex situ* conservation gap analyses of crop wild relatives from Malawi. *Genet. Resour. Crop Evol.* 68 (2), 759–771. doi: 10.1007/s10722-020-01021-3
- Khan, A. W., Garg, V., Roorkiwal, M., Golicz, A. A., Edwards, D., and Varshney, R. K. (2020). Super-pangenome by integrating the wild side of a species for accelerated crop improvement. *Trends Plant Sci.* 25 (2), 148–158. doi: 10.1016/j.tplants.2019.10.012
- Khazaei, H., Street, K., Bari, A., Mackay, M., and Stoddard, F. L. (2013). The FIGS (Focused identification of germplasm strategy) approach identifies traits related to drought adaptation in vicia faba genetic resources. *PLoS One* 8 (5), e63107. doi: 10.1371/journal.pone.0063107
- Khoury, C. K., Brush, S., Costich, D. E., Curry, H. A., de Haan, S., Engels, J. M. M., et al. (2022). Crop genetic erosion: understanding and responding to loss of crop diversity. *New Phytol.* 233 (1), 84–118. doi: 10.1111/nph.17733
- Khoury Colin, K., Carver, D., Greene Stephanie, L., Williams Karen, A., Achicanoy Harold, A., Schori, M., et al. (2020). Crop wild relatives of the united states require urgent conservation action. *Proc. Natl. Acad. Sci.* 117 (52), 33351–33357. doi: 10.1073/pnas.2007029117
- Khoury, C. K., Greene, S. L., Krishnan, S., Miller, A. J., and Moreau, T. (2019). A road map for conservation, use, and public engagement around north america's crop wild relatives and wild utilized plants. *Crop Sci.* 59 (6), 2302–2307. doi: 10.2135/cropsci2019.05.0309
- Khoury, C. K., Greene, S., Wiersema, J., Maxted, N., Jarvis, A., and Struik, P. C. (2013). An inventory of crop wild relatives of the united states. *Crop Sci.* 53 (4), 1496–1508. doi: 10.2135/cropsci2012.10.0585
- Klein, R. R., Miller, F. R., Bean, S., and Klein, P. E. (2016). Registration of 40 converted germplasm sources from the reinstated sorghum conversion program. *J. Plant Registrations* 10 (1), 57–61. doi: 10.3198/jpr2015.05.0034crg
- Koinange, E. M., Singh, S. P., and Gepts, P. (1996). Genetic control of the domestication syndrome in common bean. *Crop Sci.* 36 (4), 1037–1045. doi: 10.2135/cropsci1996.0011183X003600040037x
- Konon, O. N., D'Hont, A., Baudoin, J. P., and Mergeai, G. (2007). Cytogenetics of a new trispecies hybrid in cotton: [(*Gossypium hirsutum* L. × *G. thurberi* Tod.)₂ × *G. longicalyx* Hutch. & Lee]. *Plant Breed.* 126 (2), 176–181. doi: 10.1111/j.1439-0523.2007.01325.x
- Kouassi, A. B., Kouassi, K. B. A., Sylla, Z., Plazas, M., Fonseca, R. M., Kouassi, A., et al. (2021). Genetic parameters of drought tolerance for agronomical traits in eggplant, wild relatives, and interspecific hybrids. *Crop Sci.* 61 (1), 55–68. doi: 10.1002/csc2.20250
- Lala, S., Amri, A., and Maxted, N. (2018). Towards the conservation of crop wild relative diversity in north Africa: checklist, prioritisation and inventory. *Genet. Resour. Crop Evol.* 65 (1), 113–124. doi: 10.1007/s10722-017-0513-5
- Lam, H.-M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F.-L., et al. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat. Genet.* 42 (12), 1053–1059. doi: 10.1038/ng.715
- Landucci, F., Panella, L., Lucarini, D., Gigante, D., Donnini, D., Kell, S., et al. (2014). A prioritized inventory of crop wild relatives and wild harvested plants of Italy. *Crop Sci.* 54 (4), 1628–1644. doi: 10.2135/cropsci2013.05.0355
- Lemmon, Z. H., Reem, N. T., Dalrymple, J., Soyk, S., Swartwood, K. E., Rodriguez-Leal, D., et al. (2018). Rapid improvement of domestication traits in an orphan crop by genome editing. *Nat. Plants* 4, 766–770. doi: 10.1038/s41477-018-0259-x
- Lin, X., Armstrong, M., Baker, K., Wouters, D., Visser, R. G. F., Wolters, P. J., et al. (2020). *RLP/K* enrichment sequencing: a novel method to identify receptor-like protein (*RLP*) and receptor-like kinase (*RLK*) genes. *New Phytol.* 277 (4), 1264–1276. doi: 10.1111/nph.16608
- Lin, C.-S., Hsu, C.-T., Yuan, Y.-H., Zheng, P.-X., Wu, F.-H., Cheng, Q.-W., et al. (2022). DNA-Free CRISPR-Cas9 gene editing of wild tetraploid tomato *Solanum peruvianum* using protoplast regeneration. *Plant Physiol.* 188 (4), 1917–1930. doi: 10.1093/plphys/kiac022
- Liu, Y., Du, H., Li, P., Shen, Y., Peng, H., Liu, S., et al. (2020). Pan-genome of wild and cultivated soybeans. *Cell* 182 (1), 162–176.e113. doi: 10.1016/j.cell.2020.05.023
- Liu, Y., Wang, D., He, F., Wang, J., Joshi, T., and Xu, D. (2019). Phenotype prediction and genome-wide association study using deep convolutional neural network of soybean. *Front. Genet.* 10. doi: 10.3389/fgenet.2019.01091
- Louette, D., Charrier, A., and Berthaud, J. (1997). *In situ* conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community. *Economic Bot.* 51 (1), 20–38. doi: 10.1007/BF02910401
- Mace, E. S., Cruickshank, A. W., Tao, Y., Hunt, C. H., and Jordan, D. R. (2021). A global resource for exploring and exploiting genetic variation in sorghum crop wild relatives. *Crop Sci.* 61 (1), 150–162. doi: 10.1002/csc2.20332
- Ma, W., Qiu, Z., Song, J., Li, J., Cheng, Q., Zhai, J., et al. (2018). A deep convolutional neural network approach for predicting phenotypes from genotypes. *PLANTA* 248 (5), 1307–1318. doi: 10.1007/s00425-018-2976-9
- Mason, A. S., and Batley, J. (2015). Creating new interspecific hybrid and polyploid crops. *Trends Biotechnol.* 33 (8), 436–441. doi: 10.1016/j.tibtech.2015.06.004
- Mason, A. S., Zhang, J., Tollenaere, R., Vasquez Teuber, P., Dalton-Morgan, J., Hu, L., et al. (2015). High-throughput genotyping for species identification and diversity assessment in germplasm collections. *Mol. Ecol. Resour.* 15 (5), 1091–1101. doi: 10.1111/1755-0998.12379
- Maxted, N. (2008). *Crop wild relative conservation and use* (Wallingford, UK: CABI).
- Maxted, N., Ford-Lloyd, B. V., Jury, S., Kell, S., and Scholten, M. (2006). Towards a definition of a crop wild relative. *Biodiversity Conserv.* 15 (8), 2673–2685. doi: 10.1007/s10531-005-5409-6
- Maxted, N., Scholten, M., Codd, R., and Ford-Lloyd, B. (2007). Creation and use of a national inventory of crop wild relatives. *Biol. Conserv.* 140 (1-2), 142–159. doi: 10.1016/j.biocon.2007.08.006

- Maxted, N., and Vincent, H. (2021). Review of congruence between global crop wild relative hotspots and centres of crop origin/diversity. *Genet. Resour. Crop Evol.* 68 (4), 1283–1297. doi: 10.1007/s10722-021-01114-7
- Mertens, A., Swennen, R., Rønsted, N., Vandeloek, F., Panis, B., Sachter-Smith, G., et al. (2021). Conservation status assessment of banana crop wild relatives using species distribution modelling. *Diversity Distributions* 27 (4), 729–746. doi: 10.1111/ddi.13233
- Metwally, E., Sharshar, M., Masoud, A., Kilian, B., Sharma, S., Masry, A., et al. (2021). Development of high yielding cowpea [*Vigna unguiculata* (L.) walp.] lines with improved quality seeds through mutation and pedigree selection methods. *Horticulturae* 7 (9), 271. doi: 10.3390/horticulturae7090271
- Meyer, A., and Barton, N. (2019). Botanic Gardens Are Important Contributors to Crop Wild Relative Preservation. *Crop Sci.* 59, 2404–12. doi: 10.2135/cropsci2019.06.0358
- Mittler, R., and Blumwald, E. (2010). Genetic engineering for modern agriculture: challenges and perspectives. *Annu. Rev. Plant Biol.* 61, 443–462. doi: 10.1146/annurev-arplant-042809-112116
- Moore, J. D., Kell, S. P., Iriondo, J. M., Ford-Lloyd, B. V., and Maxted, N. (2008). CWRML: representing crop wild relative conservation and use data in XML. *BMC Bioinf.* 9 (1), 1–7. doi: 10.1186/1471-2105-9-116
- Mounce, R., Smith, P., and Brockington, S. (2017). Ex situ conservation of plant diversity in the world's botanic gardens. *Nat. Plants* 3 (10), 795–802. doi: 10.1038/s41477-017-0019-3
- Nagatoshi, Y., and Fujita, Y. (2019). Accelerating soybean breeding in a CO₂-supplemented growth chamber. *Plant Cell Physiol.* 60 (1), 77–84. doi: 10.1093/pcp/pcy189
- Nduche, M., Brehm, J. M., Abberton, M., Omosun, G., and Maxted, N. (2021). “West African Crop wild relative checklist, prioritization and inventory,” in *Genetic resources* 2(4), 55–65. doi: 10.46265/genresj.EIFL1323
- Nelson, R., Wiesner-Hanks, T., Wissler, R., and Balint-Kurti, P. (2018). Navigating complexity to breed disease-resistant crops. *Nat. Rev. Genet.* 19 (1), 21–33. doi: 10.1038/nrg.2017.82
- Nguni, D., Munkombwe, G., Mwila, G., Gaisberger, H., Brehm, J. M., Maxted, N., et al. (2019). Spatial analyses of occurrence data of crop wild relatives (CWR) taxa as tools for selection of sites for conservation of priority CWR in Zambia. *Plant Genet. Resour.* 17 (2), 103–114. doi: 10.1017/S1479262118000497
- Obsie, E. Y., Qu, H., and Drummond, F. (2020). Wild blueberry yield prediction using a combination of computer simulation and machine learning algorithms. *Comput. Electron. Agric.* 178, 105778. doi: 10.1016/j.compag.2020.105778
- Papa, R., and Gepts, P. (2003). Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from mesoamerica. *Theor. Appl. Genet.* 106 (2), 239–250. doi: 10.1007/s00122-002-1085-z
- Parker, T. A., Berny Mier y Teran, J. C., Palkovic, A., Jernstedt, J., and Gepts, P. (2020). Pod indehiscence is a domestication and aridity resilience trait in common bean. *New Phytol.* 225 (1), 558–570. doi: 10.1111/nph.16164
- Periyannan, S., Moore, J., Ayliffe, M., Bansal, U., Wang, X., Huang, L., et al. (2013). The gene Sr33, an ortholog of barley mla genes, encodes resistance to wheat stem rust race Ug99. *Science* 341 (6147), 786–788. doi: 10.1126/science.1239028
- Perrino, E. V., and Perrino, P. (2020). Crop wild relatives: know how past and present to improve future research, conservation and utilization strategies, especially in Italy: a review. *Genet. Resour. Crop Evol.* 67 (5), 1067–1105. doi: 10.1007/s10722-020-00930-7
- Perrino, E. V., and Wagensommer, R. P. (2022). Crop wild relatives (CWRs) threatened and endemic to Italy: Urgent actions for protection and use. *Biol. (Basel)* 11 (2), 193. doi: 10.3390/biology11020193
- Phillips, J., Asdal, Å., Magos Brehm, J., Rasmussen, M., and Maxted, N. (2016). *In situ* and ex situ diversity analysis of priority crop wild relatives in Norway. *Diversity Distributions* 22 (11), 1112–1126. doi: 10.1111/ddi.12470
- Phillips, J., Kyrtatzis, A., Christoudoulou, C., Kell, S., and Maxted, N. (2014). Development of a national crop wild relative conservation strategy for Cyprus. *Genet. Resour. Crop Evol.* 61 (4), 817–827. doi: 10.1007/s10722-013-0076-z
- Pimentel, D., Wilson, C., McCullum, C., Huang, R., Dwen, P., Flack, J., et al. (1997). Economic and environmental benefits of biodiversity. *BioScience* 47 (11), 747–757. doi: 10.2307/1313097
- Pironon, S., Borrell, J. S., Ondo, I., Douglas, R., Phillips, C., Khoury, C. K., et al. (2020). Toward unifying global hotspots of wild and domesticated biodiversity. *Plants* 9 (9):1128. doi: 10.3390/plants9091128
- PolicyReport (2016) *In situ and ex situ conservation, two sides of the same coin*. Available at: <https://www.cwrdiversity.org/wp/wp-content/uploads/2016/11/In-Situ-Ex-Situ-Policy-Brief.pdf>.
- Portwood, J. L., Woodhouse, M. R., Cannon, E. K., Gardiner, J. M., Harper, L. C., Schaeffer, M. L., et al. (2019). MaizeGDB 2018: the maize multi-genome genetics and genomics database. *Nucleic Acids Res.* 47 (D1), D1146–D1154.
- Postman, J., Hummer, K., Ayala-Silva, T., Bretting, P., Franko, T., Kinard, G., et al. (2010). GRIN-Global: An international project to develop a global plant genebank information management system. *Acta Hort.* 859, 49–55. doi: 10.17660/ActaHortic.2010.859.4
- Rahman, W., Brehm, J. M., Maxted, N., Phillips, J., Contreras-Toledo, A. R., Faraji, M., et al. (2021). Gap analyses of priority wild relatives of food crop in current ex situ and *in situ* conservation in Indonesia. *Biodiversity Conserv.* 30 (10), 2827–2855. doi: 10.1007/s10531-021-02225-4
- Rana, M. M., Takamatsu, T., Baslam, M., Kaneko, K., Itoh, K., Harada, N., et al. (2019). Salt tolerance improvement in rice through efficient SNP marker-assisted selection coupled with speed-breeding. *Int. J. Mol. Sci.* 20 (10), 2585–2585. doi: 10.3390/ijms20102585
- Ranganathan, J., Vennard, D., Waite, R., Dumas, P., Lipinski, B., and Searchinger, T. (2016). Shifting diets for a sustainable food future. *World Resour. Institute: Washington DC U.S.A.*
- Ran, Y., Liang, Z., and Gao, C. (2017). Current and future editing reagent delivery systems for plant genome editing. *Sci. China Life Sci.* 60 (5), 490–505. doi: 10.1007/s11427-017-9022-1
- Ratnayake, S. S., Kariyawasam, C. S., Kumar, L., Hunter, D., and Liyanage, A. S. U. (2021). Potential distribution of crop wild relatives under climate change in Sri Lanka: implications for conservation of agricultural biodiversity. *Curr. Res. Environ. Sustainability* 3, 100092. doi: 10.1016/j.crsust.2021.100092
- Raubach, S., Kilian, B., Dreher, K., Amri, A., Bassi, F. M., Boukar, O., et al. (2021). From bits to bites: Advancement of the germinate platform to support prebreeding informatics for crop wild relatives. *Crop Sci.* 61 (3), 1538–1566. doi: 10.1002/csc.2.20248
- Raza, A., Razzaq, A., Mehmood, S. S., Zou, X., Zhang, X., Lv, Y., et al. (2019). Impact of climate change on crops adaptation and strategies to tackle its outcome: A review. *Plants* 8 (2), 34. doi: 10.3390/plants8020034
- Rizal, G., Karki, S., Alcasid, M., Montecillo, F., Acebron, K., Larazo, N., et al. (2014). Shortening the breeding cycle of sorghum, a model crop for research. *Crop Sci.* 54, 520–529. doi: 10.2135/cropsci2013.07.0471
- Robinson, A., Bell, A., Dinghe, N., and Stelly, D. (2004). “Status report on introgression of reniform nematode resistance from gossypium longicalyx.” In: *Proceedings of the Beltwide Cotton Conferences*, San Antonio, Texas.
- Rosenow, D. T., Dahlberg, J. A., Stephens, J. C., Miller, F. R., Barnes, D. K., Peterson, G. C., et al. (1997). Registration of 63 converted sorghum germplasm lines from the sorghum conversion program. *Crop Sci.* 37 (4), 1399–1400. doi: 10.2135/cropsci1997.0011183X003700040090x
- Saintenac, C., Zhang, W., Salcedo, A., Rouse Matthew, N., Trick Harold, N., Akhunov, E., et al. (2013). Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. *Science* 341 (6147), 783–786. doi: 10.1126/science.1239022
- Sansaloni, C., Franco, J., Santos, B., Percival-Alwyn, L., Singh, S., Petroli, C., et al. (2020). Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. *Nat. Commun.* 11 (1), 4572. doi: 10.1038/s41467-020-18404-w
- Saxena, R. K., Edwards, D., and Varshney, R. K. (2014). Structural variations in plant genomes. *Briefings Funct. Genomics* 13 (4), 296–307. doi: 10.1093/bfpg/elu016
- Schouten, H. J., Tikunov, Y., Verkerke, W., Finkers, R., Bovy, A., Bai, Y., et al. (2019). Breeding has increased the diversity of cultivated tomato in the Netherlands. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01606
- Schwacke, R., Ponce-Soto, G. Y., Krause, K., Bolger, A. M., Arsova, B., Hallab, A., et al. (2019). MapMan4: a refined protein classification and annotation framework applicable to multi-omics data analysis. *Mol. Plant* 12 (6), 879–892. doi: 10.1016/j.molp.2019.01.003
- Shan-e-Ali Zaidi, S., and Mansoor, S. (2017). Viral vectors for plant genome engineering. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00539/bibtex
- Shaw, P. D., Raubach, S., Hearne, S. J., Dreher, K., Bryan, G., McKenzie, G., et al. (2017). Germinate 3: development of a common platform to support the distribution of experimental data on crop wild relatives. *Crop Sci.* 57 (3), 1259–1273. doi: 10.2135/cropsci2016.09.0814
- Song, J.-M., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., et al. (2020). Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of brassica napus. *Nat. Plants* 6 (1), 34–45. doi: 10.1038/s41477-019-0577-7
- Stephens, J. C., Miller, F. R., and Rosenow, D. T. (1967). Conversion of alien sorghums to early combine Genotypes1. *Crop Sci.* 7 (4), 396–396. doi: 10.2135/cropsci1967.0011183X000700040036x
- Steuernagel, B., Periyannan, S. K., Hernández-Pinzón, I., Witek, K., Rouse, M. N., Yu, G., et al. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* 34 (6), 652–655. doi: 10.1038/nbt.3543
- Stoilova, T., van Zonneveld, M., Roothaert, R., and Schreinemachers, P. (2019). Connecting genebanks to farmers in East Africa through the distribution of

- vegetable seed kits. *Plant Genet. Resources: Characterization Utilization* 17 (3), 306–309. doi: 10.1017/S1479262119000017
- Summers, R. W., and Brown, J. K. M. (2013). Constraints on breeding for disease resistance in commercially competitive wheat cultivars. *Plant Pathol.* 62 (S1), 115–121. doi: 10.1111/ppa.12165
- Tao, Y., Luo, H., Xu, J., Cruickshank, A., Zhao, X., Teng, F., et al. (2021). Extensive variation within the pan-genome of cultivated and wild sorghum. *Nat. Plants* 7 (6), 766–773. doi: 10.1038/s41477-021-00925-x
- Tas, N., West, G., Kircalioglu, G., Topaloglu, S. B., Phillips, J., Kell, S., et al. (2019). Conservation gap analysis of crop wild relatives in Turkey. *Plant Genet. Resour.* 17 (2), 164–173. doi: 10.1017/S1479262118000564
- Tay Fernandez, C. G., Nestor, B. J., Danilovic, M. F., Gill, M., Peterit, J., Bayer, P. E., et al. (2022). Pangenomes as a resource to accelerate breeding of underutilised crop species. *Int. J. Mol. Sci.* 23 (5), 2671. doi: 10.3390/ijms23052671
- Taylor, N., Holubec, V., Chobot, K., Parra-Quijano, M., Maxted, N., and Kell, S. (2013). Systematic crop wild relative conservation planning for the Czech republic. *Crop Wild relative* 9, 5–9. doi: 10.1079/9781845930998.000
- Teso, M. L. R., Lamas, E. T., Parra-Quijano, M., de la Rosa, L., Fajardo, J., and Iriondo, J. M. (2018). National inventory and prioritization of crop wild relatives in Spain. *Genet. Resour. Crop Evol.* 65 (4), 1237–1253. doi: 10.1007/s10722-018-0610-0
- Tyack, N., Dempewolf, H., and Khoury, C. K. (2020). The potential of payment for ecosystem services for crop wild relative conservation. *Plants* 9 (10), 1305. doi: 10.3390/plants9101305
- Van Bel, M., Silvestri, F., Weitz, E. M., Krefth, L., Botzki, A., Coppens, F., et al. (2022). PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. *Nucleic Acids Res.* 50 (D1), D1468–D1474. doi: 10.1093/nar/gkab1024
- van Bemmelen van der Plaat, A., van Treuren, R., and van Hintum, T. J. L. (2021). Reliable genomic strategies for species classification of plant genetic resources. *BMC Bioinf.* 22 (1), 173. doi: 10.1186/s12859-021-04018-6
- Vavilov, N. (1926). Center of origin of cultivated plants. *Papers Appl. Botany Genet. Plant Breeding* 16, 1–248.
- Vavilov, N. I., Vavilov, M. I., and Dorofeev, V. F. (1992). *Origin and geography of cultivated plants* (Cambridge: Cambridge University Press).
- Vendelbo, N. M., Mahmood, K., Steuernagel, B., Wulff, B. B. H., Sarup, P., Hommoller, M. S., et al. (2022). Discovery of resistance genes in rye by targeted long-read sequencing and association genetics. *Cells* 11 (8). doi: 10.3390/cells11081273
- Vincent, H., Wiersema, J., Kell, S., Fielder, H., Dobbie, S., Castañeda-Álvarez, N. P., et al. (2013). A prioritized crop wild relative inventory to help underpin global food security. *Biol. Conserv.* 167, 265–275. doi: 10.1016/j.biocon.2013.08.011
- Vincent, H., Hole, D., and Maxted, N. (2022). Congruence between global crop wild relative hotspots and biodiversity hotspots. *Biological Conservation* 265, 109432. doi: 10.1016/j.biocon.2021.109432
- Viruel, J., Kantar, M. B., Gargiulo, R., Hesketh-Prichard, P., Leong, N., Cockel, C., et al. (2021). Crop wild phylorelatives (CWPs): phylogenetic distance, cytogenetic compatibility and breeding system data enable estimation of crop wild relative gene pool classification. *Botanical J. Linn. Soc.* 195 (1), 1–33. doi: 10.1093/botlinnean/boaa064
- Wambugu, P. W., Ndjondjop, M.-N., and Henry, R. J. (2018). Role of genomics in promoting the utilization of plant genetic resources in genebanks. *Briefings Funct. Genomics* 17 (3), 198–206. doi: 10.1093/bfpp/ely014
- Wang, M., Yang, J., Wan, J., Tao, D., Zhou, J., Yu, D., et al. (2020). A hybrid sterile locus leads to the linkage drag of interspecific hybrid progenies. *Plant Divers.* 42 (5), 370–375. doi: 10.1016/j.pld.2020.07.003
- Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M.-D., et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. Plants* 4, 23–29. doi: 10.1038/s41477-017-0083-8
- Wegrzyn, J. L., Liechty, J. D., Stevens, K. A., Wu, L.-S., Loopstra, C. A., Vasquez-Gross, H. A., et al. (2014). Unique features of the loblolly pine (*Pinus taeda* L.) megagenome revealed through sequence annotation. *Genetics* 196 (3), 891–909. doi: 10.1534/genetics.113.159996
- Wilkinson, P. A., Allen, A. M., Tyrrell, S., Wingen, L. U., Bian, X., Winfield, M. O., et al. (2020). CerealsDB—new tools for the analysis of the wheat genome: update 2020. *Database* 2020, 1–13. doi: 10.1093/database/baaa060
- Witek, K., Jupe, F., Witek, A. I., Baker, D., Clark, M. D., and Jones, J. D. G. (2016). Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nat. Biotechnol.* 34 (6), 656–660. doi: 10.1038/nbt.3540
- Xiang, Z., Chen, Y., Chen, Y., Zhang, L., Liu, M., Mao, D., et al. (2022). Agrobacterium-mediated high-efficiency genetic transformation and genome editing of chaling common wild rice (*Oryza rufipogon* griff.) using scutellum tissue of embryos in mature seeds. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.849666
- Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F., Li, X., et al. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30 (1), 105–111. doi: 10.1038/nbt.2050
- Yahiaoui, N., Kaur, N., and Keller, B. (2009). Independent evolution of functional Pm3 resistance genes in wild tetraploid wheat and domesticated bread wheat. *Plant J.* 57 (5), 846–856. doi: 10.1111/j.1365-3113.2008.03731.x
- Yik, C. P., and Birchfield, W. (1984). Resistant germplasm in gossypium species and related plants to rotylechulus reniformis. *J. Nematol* 16 (2), 146–153.
- Yocca, A. E., and Edger, P. P. (2022). Machine learning approaches to identify core and dispensable genes in pangenomes. *Plant Genome* 15 (1), e20135. doi: 10.1002/tpg2.20135
- Yu, H., Lin, T., Meng, X., Du, H., Zhang, J., Liu, G., et al. (2021). A route to *de novo* domestication of wild allotetraploid rice. *Cell* 184 (5), 1156–1170. doi: 10.1016/j.cell.2021.01.013
- Zair, W., Maxted, N., and Amri, A. (2018). Setting conservation priorities for crop wild relatives in the fertile crescent. *Genet. Resour. Crop Evol.* 65 (3), 855–863. doi: 10.1007/s10722-017-0576-3
- Zair, W., Maxted, N., Brehm, J. M., and Amri, A. (2021). Ex situ and *in situ* conservation gap analysis of crop wild relative diversity in the fertile crescent of the middle East. *Genet. Resour. Crop Evol.* 68 (2), 693–709. doi: 10.1007/s10722-020-01017-z
- Zhang, H., Demirer, G. S., Zhang, H., Ye, T., Goh, N. S., Aditham, A. J., et al. (2019). DNA Nanostructures coordinate gene silencing in mature plants. *Proc. Natl. Acad. Sci.* 116 (15), 7543–7548. doi: 10.1073/pnas.1818290116
- Zhan, X., Lu, Y., Zhu, J. K., and Botella, J. R. (2021). Genome editing for plant research and crop improvement. *J. Integr. Plant Biol.* 63 (1), 3–33. doi: 10.1111/jipb.13063
- Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., Li, W., et al. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33 (4), 408–414. doi: 10.1038/nbt.3096
- Zhu, H., Li, C., and Gao, C. (2020). Applications of CRISPR-cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* 21, 661–677. doi: 10.1038/s41580-020-00288-9
- Zsögön, A., Čermák, T., Naves, E. R., Notini, M. M., Edel, K. H., Weinl, S., et al. (2018). *De novo* domestication of wild tomato using genome editing. *Nat. Biotechnol.* 36 (12), 1211–1216. doi: 10.1038/nbt.4272