The role of genetics in mainstreaming the production of new and orphan crops to diversify food systems and support human nutrition

Citation for published version:
Dawson, IK, Powell, W, Hendre, P, Bani, J, Hickey, JM, Kindt, R, Hoad, S, Hale, I & Jamnadass, R 2019, 'The role of genetics in mainstreaming the production of new and orphan crops to diversify food systems and support human nutrition', *New Phytologist*. https://doi.org/10.1111/nph.15895

Digital Object Identifier (DOI):
10.1111/nph.15895

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
New Phytologist

DR IAN K DAWSON (Orcid ID : 0000-0003-4245-6973)

PROF. WAYNE POWELL (Orcid ID : 0000-0002-5612-3398)

Article type : Commissioned Material - Tansley Review

The role of genetics in mainstreaming the production of new and orphan crops to diversify food systems and support human nutrition

Ian K Dawson[1,2], Wayne Powell[1], Prasad Hendre[2], Jon Bančič[1,3], John M Hickey[3], Roeland Kindt[2], Steve Hoad[1], Iago Hale[4], Ramni Jamnadass[2]

[1] Scotland's Rural College (SRUC), Kings Buildings, West Mains Road, Edinburgh, EH9 3JG, Scotland

[2] World Agroforestry (ICRAF), Headquarters, PO Box 30677, Nairobi, Kenya

[3] The Roslin Institute, Easter Bush Campus, University of Edinburgh, Midlothian, EH25 9RG, Scotland

[4] University of New Hampshire, Durham, NH 03824, USA

Author for correspondence: Ian K Dawson

Tel. +44 1904 628 367

Email: iankdawson@aol.com

Received: 15 February 2019

Accepted: 28 April 2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/nph.15895

ORCID IDs (where available):

Ian Dawson https://orcid.org/0000-0003-1345-3642

Wayne Powell https://orcid.org/0000-0002-5612-3398

Prasad Hendre https://orcid.org/0000-0003-1691-183X

Roeland Kindt https://orcid.org/0000-0002-7672-0712

Steve Hoad https://orcid.org/0000-0003-3621-4265

Iago Hale https://orcid.org/0000-0002-8558-7861

Ramni Jamnadass https://orcid.org/0000-0003-2416-8361

Contents

Summary

Especially in low income nations, new and orphan crops provide important opportunities to improve diet quality and the sustainability of food production, being rich in nutrients, capable of fitting into multiple niches in production systems, and relatively adapted to low input conditions. The evolving space for these crops in production systems presents particular genetic improvement requirements that extensive gene pools are able to accommodate. Particular needs for genetic development identified in part with plant breeders relate to three areas of fundamental importance for addressing food production and human demographic trends and associated challenges, which are: facilitating integration into production systems; improving the processability of crop products; and reducing farm labour requirements. Here, we relate diverse involved target genes and crop development techniques. These techniques include transgressive methods that involve defining exemplar crop models for effective new and orphan crop improvement pathways. Research on new and orphan crops not only supports the genetic improvement of these crops, but they serve as important models for understanding crop evolutionary processes more broadly, guiding further major crop evolution. The bridging position of orphan crops between new and major crops provides unique opportunities for investigating genetic approaches for *de novo* domestications and major crop 'rewildings'.

Keywords: breeding approaches, crop harvestability, crop integration, crop processability, model crop exemplars, orthologous genes.

I. Introduction

Global food production has homogenised as an ever-narrower range of calorie-rich but nutritionally-limited and resource-intensive crops has increased in dominance (Khoury *et al*., 2014). This has enhanced energy availability in diets but endangers human and environmental health by contributing to hidden hunger (von Grebmer *et al*., 2014), climate-related food production shocks (Global Food Security, 2015) and planetary resource depletion (Rockström *et al*., 2009). Rediversifying crop production is important to promote a wider range of healthier foods and more sustainable and stable production systems (von Grebmer *et al*., 2014). Efforts to promote diversification however require that policies and research priorities change (Khoury & Jarvis, 2014; Gillespie & van den Bold, 2017; Willett *et al*., 2019). In a revised agenda, new and orphan crops rich in valuable micro- and macro-nutrients, capable of fitting into multiple niches in production systems (where they provide environmental services as well as direct provisioning services), and relatively adapted to low input conditions, have an important role (Gruber, 2017; AOCC, 2019; Mustafa *et al*., 2019). These plants either are harvested at present from the wild and are candidates for cultivation (potential new crops) or have already entered the domestication process but are generally only grown and valued locally or regionally (orphan crops); in both cases, only limited production research is currently underway on them (Dawson *et al*., 2018b). Despite this neglect, new and orphan crops have received increased media attention recently (e.g., Economist, 2017), based on their potential to address multiple UN Sustainable Development Goals (UN, 2019) in the low income nations of Africa (AOCC, 2019), Asia (CFF, 2019) and Latin America (LATINCROP, 2019), and due to Western consumers' interests in new, healthier foods.

As well as new and orphan crops' production values, they are attractive candidates for research by biologists. This is because crop domestication – defined broadly as the genetic changes involved in bringing a crop into cultivation and in its continued development within agriculture – has long been recognised as providing fascinating insights into fundamental evolutionary processes (Darwin, 1859). Clearly, new, orphan and major crops position on a domestication continuum. Furthermore, the range of locations over which their extant genetic diversity is distributed also varies. Together, these points mean that a spectrum of these plants can allow the extrapolation domains of various possible genetic improvement approaches to be explored, as outlined in Figure 1.

In addition to their practical and research values, the technical environment for undertaking new and orphan crop genetic studies has greatly improved in the last decade due to significant cost reductions in genome characterisation approaches (Varshney *et al*., 2012). Along with advances in 'speed breeding' (Watson *et al*., 2018), in participatory improvement methods (Weltzien & Christinck, 2017) and in statistical approaches that support crop development (e.g., Meuwissen *et al*., 2001; Lasky *et al*., 2018), the landscape for new and orphan crop genetical study has therefore recently been transformed.

The above observations all argue for more attention to be paid to new and orphan crops. Here, we consider how the production of these crops may support human and environmental health objectives, paying particular attention to the situation in tropical and subtropical low income nations. Deficits in key dietary nutrients are often high in these countries, but they also often contain extensive inter- and intra-specific variation in wild and agricultural plants that could be better utilised for biodiversity-based, sustainable food solutions (Jamnadass *et al*., 2011). Below, we first address the context of global crop production, considering trends over the last half century that inform possible new and orphan crop genetic interventions for initial or wider integration of these plants into agriculture. We then consider genetic improvement objectives, drawing on existing knowledge of the crop 'domestication syndrome' (Meyer *et al*., 2012), our own analysis of plant breeders' perspectives on crop development needs, and considering other food system stakeholders' requirements. We particularly focus on traits and examples of underlying genes to address food production and human demographic trends and associated challenges in three areas of fundamental importance: to support the integration of crops into production systems; to increase crop product processability; and to reduce the farm labour requirements of production. We then relate approaches for the genetic improvement of new and orphan crops, considering the role of orthologous gene sequences in trait evolution. As part of this exercise, we illustrate an approach for defining appropriate genetic improvement pathways for a range of exemplar-requiring new and orphan crops, based on comparison with a panel of more widely understood crop models.

Our intention through this review is to indicate genetics-based research avenues to support the mainstreaming of new and orphan crops in food production systems. In addition, we illustrate how research on these plants can contribute to major crop evolution. Clearly, crop improvement is only one aspect to be addressed in mainstreaming production. Further interdisciplinary work, such as to understand the social and economic drivers of consumer demand, is also required (Dawson *et al*., 2018b). In addition, policy issues related to the use of genetic technologies, such as the effective application of the Nagoya Protocol (on access to, and the fair and equitable sharing of the benefits arising from the utilisation of, genetic resources), also need to be addressed for new and orphan crops (e.g., Østerberg *et al*., 2017; Halewood *et al*., 2018). These aspects, while outside the scope of the current review, are also of high importance.

II. Trends in crop production that inform new and orphan crop promotion

In their analysis of global crop production trends, Khoury *et al*. (2014) identified crops that were relative over-performers ('winners') and under-performers ('losers') over the approximate half century of 1961 to 2009 in terms of total food supply. To explore the wider production characteristics of winner and loser crops, we further analysed a representative selection of them. Our analysis (Fig. 2) indicated that winner crops are more likely to be produced in lower diversity production systems (tending to monoculture) than are loser crops. This is consistent with the global reductions in farm production system heterogeneity over recent decades that have been explored by other authors (e.g., Clay, 2004; Donald, 2004). Current global production trends thus not only result in lower crop food diversity, endangering humans' nutritional security, but call into question the continued availability of agrobiodiversity-related environmental services within farm landscapes and therefore the sustainability of food production more broadly (Cardinale *et al*., 2012). Designing new and orphan crops to better support the maintenance and development of diverse production systems is therefore a doubly crucial objective. In these systems, the intention should be that new and orphan crops complement the production of other crops rather than simply substitute for them, requiring appropriate spatial and temporal integration (Dawson *et al*., 2018a).

To help determine where investments in productivity improvements that are generally considered a fundamental requirement in plant breeding could drive production diversification with new and orphan crops, we again reviewed available crop production data. We assessed the relative contributions of yield (production per unit area, with genetic and/or agronomic contributions possible) and total area planted to changes in global crop output over the last half century for a panel of 35 exemplar crops (Fig. 3 and Supporting Information Notes S1). Our analysis identified a group of nine case study crops where yield contributions to changes in output appeared markedly lower than the established trend line. A comparison of these crops with the wider panel revealed that most had relatively low annual global gross production values in monetary terms, a situation equating to the majority of new and orphan crops that are used locally and regionally only.

It seems reasonable to assume that monetary production value is a proxy for historical levels of investment in plant breeding. Thus, the observed differences in crop performance that likely relate to breeding investment in our analysis indicate the importance of new breeding efforts to support new and orphan crop development. Furthermore, such differences in performance in relation to investment highlight the potential transformative role that new, cheaper advanced breeding approaches could have in reducing investment barriers to support significant production gains (Varshney *et al*., 2012). Our analysis would suggest that investment in advanced breeding methods is of particular importance for addressing potential improvement challenges for perennial, vegetatively-propagated crops.

III. Genetic improvement objectives for new and orphan crops

III.1 New and orphan crop development and the crop domestication syndrome

The starting point for new and orphan crop development is a broad understanding of the 'domestication syndrome'. This is generally revealed by comparing crops' phenotypes with either the extant descendant generations or the archaeological remains of their wild progenitors (Meyer *et al*., 2012). Features of the syndrome commonly reported for annual crops include a reduced ability to disperse seed, more synchronous seed germination, increased seed size, reduced chemical defences and alterations in reproductive shoot architecture (Larson *et al*., 2014). The syndrome is less well defined for perennial crops (Miller & Gross, 2011) but in the case of fruit trees features include a shift from seed to vegetative propagation, increased regularity in fruit bearing, enhanced fruit size and decreased plant stature (Clement, 2004; Goldschmidt, 2013).

The genes underlying domestication have been partially determined in a range of crops. In an authoritative review, Meyer and Purugganan (2013) listed 60 genes whose variants were reported to be involved. Of these, 37 were reported to encode transcription factors (see also Schilling *et al*., 2018) and 14 to encode enzymes. Loss-of-function alleles were found to be the most common type of causative change, followed by alleles varying in cis-regulatory elements altering gene expression. Missense mutations (altering protein function) were however also not infrequent. Based on Meyer and Purugganan's (2013) compilation, loss-of-function mutations appeared more often associated with 'domestication' genes (that control the classic domestication syndrome) and positive change-of-function mutations with 'diversification' genes (which allow crops to adapt to particular uses and agro-ecological environments). Under this typology, 'domestication' genes may be initial targets for manipulation in new crop development, while ‘diversification’ genes may be targets in orphan crops that have already passed through initial crop development stages. In the rest of this review, we however generally refer to both these sets of genes as ‘domestication-related’ as the distinction between categories is not always clear or useful.

Around half of the genes compiled by Meyer and Purugganan (2013) were related to fruit and seed properties. Among these genes, those controlling composition and the palatability and processability of crop food parts were most prominent. For example, the *WAXY* gene in rice (*Oryza sativa*) (and orthologues in other crops) controls the amylose versus amylopectin ratio in grain starch, which determines grain processability (as discussed further later in this review). In addition, just over one-third of the compiled genes were reported to influence plant architecture or crop flowering time, both of which are important features for determining crop integration into production systems. Finally, seven of the identified genes were related to seed head non-shattering and thus to crop harvesting efficiency and crop labour requirements. Processability, integration into production systems and the labour requirements of production have all been identified as important features for new and orphan development, as we discuss further below. A focus on several of the genes compiled by Meyer and Purugganan (2013) is therefore of relevance.

III.2. Stakeholders’ perspectives on traits for new and orphan crop improvement

Understanding producers’ constraints is crucial for determining sound improvement objectives for new and orphan crops, but to date little systematic information has been available. To help remedy this gap, we have gathered information from plant breeders on new and orphan crop improvement targets for Sub-Saharan Africa, where these plants are recognised to have a particularly important role to play in supporting human nutrition and sustainable agriculture (AOCC, 2019). Although a survey of breeders’ views can only provide a partial picture of crop development needs since breeders are only one stakeholder group in crop promotion (along with farmers, consumers, retailers, food processors, etc.; Dawson *et al.*, 2018b), they are perhaps in the best position to grasp sector-wide concerns that can inform crop improvement targets. In addition, existing contact networks mean that they are a relatively easy stakeholder group to gather information from.

Our survey of breeders' views (described in Supporting Information Notes S2) indicated that crop pest and/or disease attack was the most frequently mentioned priority genetic or management constraint for new and orphan crops, while lack of access to suitable planting material was by far the most mentioned important input constraint (Fig. 4a), echoing concerns on varietal delivery that we return to later in this review. Consistent with these production constraints, breeders most mentioned pest and disease tolerance or resistance as the priority trait category for genetic improvement action, followed by yield *per se* (i.e., production independent of pests, diseases, etc., that also influence yield). In addition, improved harvestability was the fifth most mentioned important area for genetic improvement (Fig. 4b).

Our survey of breeders also indicated that improvement in crop planting and/or establishment methods was the most mentioned priority agronomic management intervention required to support new and orphan crop production, closely followed by soil fertilisation measures. The proper timing of seasonal field activities was the fourth most mentioned required agronomic intervention and the diversification of production systems the fifth (Fig. 4c). As expected, the priority constraints and interventions mentioned by breeders depended on the part of the plant used for food (Fig. 4d). Significantly, when asked about the likelihood of success of their suggested priority interventions, breeders considered agronomic management actions to be more likely to be successful than genetic improvement actions (Fig. 4e). They however believed both types of action to overall have high potential for success, suggesting a useful role for a variety of breeder-supported context-specific genetic improvement methods, in conjunction with agronomic developments. Below, we further consider the results of our breeders' survey in the context of additional stakeholders' constraints and the global trends that also inform the efficient production and use of new and orphan crops, under three trait categories of specific importance.

Traits for greater production integration

Breeders' emphasis on yield, along with knowledge of global crop production and consumption trends (described in Section II), support the view that diversification of the world's crop portfolio requires productivity enhancements in new and orphan crops, to enable them to successfully compete with major crops for farmers' attention (Tadele, 2017). Diversification is however not only about increasing the range of crops grown, but is concerned with developing more efficient, sustainable and stable integrated production systems through approaches such as intercropping

(Brooker *et al*., 2015). An emphasis on traits that maximise positive crop-crop interactions in terms of yields, sustainability and stability is therefore crucial for new and orphan crop development. This requirement appears not to have been fully recognised by the plant breeders included in our survey: while several breeders indicated the importance of crop diversification as an agronomic management intervention, less attention was given to this aspect in the trait categories identified for genetic improvement. This discrepancy could indicate either an inherent difficulty in intercrop breeding or a conceptual disconnect in breeders' current thinking, perhaps due to their tendency to work at any one time on only a single crop.

Insights into plant species' interactions in natural ecosystems may be useful for designing improved crop-crop interactions. Studies reporting the genetics (and epigenetics; Alonso *et al*., 2019) of reciprocal helping between plants are however currently relatively scarce; strategies have though been outlined through which natural genetic variants underlying mutualisms between pairs of plant species could be characterised (Subrahmaniam *et al*., 2018). Obvious 'interaction traits', likely to influence resource-use complementarity or conflict among crops, are those related to plant architecture, growth rate, mycorrhizal associations and phenology (Vandermeer, 1992; Litrico & Violle, 2015). The currently cultivated gene pools of orphan crops still contain variation in important interaction traits because this diversity has not been lost through monoculture breeding as for the advanced cultivars of major crops (Francis & Smith, 1985); there are therefore significant opportunities for designing more effective intercrop systems involving them. This depends of course on suitable breeding methods being made available, a topic we return to below.

Traits for increased product processability

Our current survey of production constraints only obtained information from plant breeders, but as already noted it is also important to consult others regarding crop target traits, including farmers, consumers, retailers and food processors, in order to 'co-construct' more optimal crop development targets. For consumers, traits related to nutritional content, food acceptability, palatability and cookability are especially important. In new and orphan root crops and legume seeds in particular, the presence of anti-nutritional compounds such as phytic acid, saponins, polyphenols, lathyrogens, α-galactosides, protease inhibitors, α-amylase inhibitors and lectins can be of concern (e.g., Sousa *et al*., 2015). Reductions in these compounds mean that foods require less cooking or other processing to remove them and make consumption safe (Yerra *et al*., 2015). In turn, this allows poor consumers to make healthier food choices. This is because the high energy costs for cooking these foods, which place a large burden on families' finances, are reduced, as are the labour requirements of food preparation, which fall especially on women (Balmer, 2007). Because a number of anti-nutritional compounds play important roles in protecting new and orphan crops from pest and disease attack, however, breeding objectives may focus on altering plant part allocation of these chemicals (e.g., avoiding the edible portion of the crop) or increasing their lability during cooking or other processing, rather than their reduction or removal *per se* (Nour-Eldin & Halkier, 2013).

With the increasing reliance by growing urban populations in low income nations on processed foods (Popkin *et al*., 2012), improvements in a range of processability traits for new and orphan crops is a priority. This allows the wider incorporation of nutritious new and orphan crop ingredients in processed food reformulations. Of relevance is the broad physical properties of ingredients and their chemical compositions, which influence flavour, texture, stability and overall consumer acceptance (e.g., Sun-Waterhouse *et al*., 2014). A good example is the ratio of amylose to amylopectin in cereal starches: this influences the functional properties of derived processed foods as well as their nutritional and physical characteristics (Lagassé *et al*., 2006). The food industry is particularly interested in identifying novel functional ingredients as surfactants, thickeners and strain-hardening biopolymers that can support more efficient, healthier processed food production; with their diverse characteristics, new and orphan crops may present novel opportunities for such uses, once they have been more fully characterised through tensiometry, rheometry and other analytical approaches to measure food properties (e.g., Bakare *et al*., 2016).

Traits for reduced farm labour requirements

The importance of reducing the labour requirements of new and orphan crop production, and of coordinating these requirements with other farm activities, is evident from our survey of breeders who indicated the need to improve crop harvestability via genetic means and the need for attention to the proper timing of seasonal field activities. The significant rural-to-urban transition currently underway in many low income nations (Kessides, 2005) reinforces the need for reducing rural labour requirements. At the same time, new concentrations of available labour in urban areas may support urban and peri-urban food transformation, reinforcing the importance of improving crop processability traits.

Trait categories influencing rural labour requirements and/or the timing of these requirements include seed and fruit dispersal or retention, seed and fruit size, plant form and crop phenology. Plant form and phenology are also crucial for crop integration, as described above, while variation in crop phenology is also important for avoiding seasonal gluts in food supply that affect market profitability and wastage. Especially for perennial crops, the length of the plant's juvenile phase is an important factor determining the return to labour. For *de novo* crops, the vision of reducing the labour needed to collect from the wild may serve as an important stimulus for initial cultivation (Schippmann, 2002), thereby implicating a need to focus on basic genetic traits related to propagation and *ex situ* establishment ability.

III.3. Candidate genes for new and orphan crop priority improvement trait categories

Some of the high priority trait categories identified for new and orphan crop improvement, such as pest and disease resistance and yield *per se*, align with the development profiles of most major crops. In this section, however, we focus on the three trait categories of specific importance for driving new and orphan crop development and adoption that were identified above. Illustrative cases of involved genes, which may present potential targets for manipulation in crop development, are summarised in Table 1. Specific examples are also given below.

For production system integration, the important feature of flowering time determines the maturity date of a crop and therefore its intercrop effectiveness (Yu *et al*., 2015). The genes controlling flowering pathways have been identified in a range of crops (Nakamichi, 2015) and the involvement of orthologous sequences across crops has been established (Calixto *et al*., 2015). Reduced photoperiod sensitivity, resulting from allelic variation in a subset of these genes, has played an enormous role in the historic range expansion of many major crops. In barley (*Hordeum vulgare*), for example, a network of ~20 circadian clock-related genes are known to modulate flowering time; the northern expansion of the crop from the Fertile Crescent was associated with the emergence of day-length insensitive forms (Russell *et al*., 2016). The manipulation of related genes in new and orphan crops could similarly facilitate range expansion and support the development of more effective intercrop combinations.

For product processability, examples related to anti-nutritional compounds and starch chemical composition are illustrative of possible genetic manipulations. In several legumes, phytates are the primary reserve of phosphorous in the seed, but these chelate iron and zinc that are essential in human diets (Petry *et al*., 2015). A low phytic acid mutant isolated in common bean (*Phaseolus vulgaris*), associated with change in an ABC transporter gene, demonstrated enhanced iron bioavailability in porridge made from its dried seed, reducing the cooking time needed to reach acceptable iron absorption levels (Petry *et al*. 2013). Various orphan crop legume seeds with high phytate levels may have levels similarly reduced through related mutations. In the case of starch composition, as already noted the ratio of amylose to amylopectin in rice grain is controlled by the granule-bound starch synthase gene *WAXY*, while its orthologues have a similar function in other cereals and pseudo-cereals, including the orphan crops of foxtail millet (*Setaria italica*) and the grain amaranths (e.g., *Amaranthus cruentus*). The waxiness of grain not only affects its attractiveness for consumers, but it influences the food processing and digestibility characteristics of seed, not always in beneficial ways for modern diets. For example, waxy grain types may be easier to process, but they may also have a higher glycaemic index that contributes to type 2 diabetes risk in humans (Kaur *et al*., 2016). Trade-offs in reaching breeding objectives for processability traits are therefore required.

Regarding labour requirements, the standard domestication syndrome trait of seed or fruit retention is crucial in influencing crop harvestability (Meyer *et al*., 2012). In addition, fruit size is an important characteristic, especially for fleshy-fruited crops where the ripe fruit is eaten whole. This is because larger fruits are easier to harvest to reach the same collected weight, particularly when the crop is handpicked. The genetic control of fruit size has been extensively researched in tomato (*Solanum lycopersicum*), the model species for other fleshy-fruited crops (van der Knaap *et al*., 2014); some of the identified genes are known to have orthologues in other plants.

IV. Approaches for genetic improvement

IV.1. Available genetic improvement methods for new and orphan crops

Ideotype targets and an understanding of the inheritance and genetic architecture of defining traits determine suitable approaches for the genetic improvement of any particular new or orphan crop (Section III). Especially in low income nations, it is also necessary to consider how improved varieties will be delivered to farmers. A detailed assessment of planting material delivery options is outside the scope of this review, but strategies are specific to breeding approach (Walker *et al*., 2014). In general, however, an emphasis on working with farmers in varietal evaluation and in the multiplication of planting stock is a useful means of building effective bridges between crop breeding and crop production (Weltzien & Christinck, 2017). This context should be considered when selecting from the possible breeding and selection options described below.

Advanced and conventional breeding

Marker-assisted selection is used widely in major crop development and has begun to be applied to orphan crops. Examples include the annual orphan crops of foxtail millet (*Setaria italica*; Jia *et al*., 2013) and pigeonpea (*Cajanus cajan*; Varshney *et al*., 2017) and, increasingly, a range of perennial plants (Iwata *et al*., 2016; Migicovsky & Myles, 2017). However, the relatively high costs of phenotyping remain a constraint in most cases (Varshney *et al*., 2012). This is especially so for perennial crops that require several years of growth before they can be properly evaluated and that have large life forms that demand considerable space in field trials.

Applying genome-wide association scans, Cichy *et al*. (2015) identified genomic regions associated with variation in the so-called “cooking time trait” in a diversity panel of common beans. Although perhaps not strictly an orphan crop itself because of relatively high research investments, common bean can be considered representative of several orphan legumes. The finding of associations between specific genomic regions and cooking time in legumes is important because long cooking times limit the more efficient utilisation of the seed as food (see discussion in Section III.2). Cichy *et al*. (2015) found statistically significant associations between cooking time and SNPs on three *P.*

vulgaris chromosomes, with the strongest associations on chromosome 6 (Pv06). Proximate coding sequences on Pv06 included two cation/H + exchanger genes, one homologous to *AtCHX3* and the other to *AtCHX4* (further information in Table 1). In arabidopsis (*Arabidopsis thaliana*) these genes are involved in calcium transport, which corresponds with evidence that Ca^{2+} plays an important role in storage-induced increases in common bean's required cooking time (Jones & Boulter, 1983). Similar to major crops (Liu & Yan, 2019), there is clearly high future potential for the further application of genome-wide association scans to dissect quantitative traits for new and orphan crops, if appropriate resources are allocated to phenotyping.

Genomic selection (Meuwissen *et al*., 2001) uses phenotypic and genomic data collected from training populations to predict the breeding value of genome-characterised but un-phenotyped breeding materials (known as genomic estimated breeding values). To date the approach has been most effectively adopted for complex trait breeding in animals (Georges *et al*., 2019), but it is increasingly being used to breed for polygenic traits in plants (Crossa *et al*., 2017), exploiting cross-sectoral synergies in possible methods (Hickey *et al*., 2017). As currently practised, the accuracy of prediction quickly decays as a function of the genetic distance between the training and experimental germplasm sets; for example, comparisons across animal breeds can be difficult (Hayes *et al*., 2009). However, advances are being made to extend useful comparisons to more distantly related materials by considering sequence context (Druet *et al*., 2014) and wider biological priors (e.g., variant annotations, candidate genes and known causal mutations; MacLeod *et al*., 2016). The ability to expand comparisons from relatively well studied crops to genetically-related but under-phenotyped new and orphan crops could be of key importance.

The efficacy of genomic selection is currently being tested on cassava (*Manihot esculenta*), a vegetatively-propagated orphan annual root crop (Wolfe *et al*., 2017). A further orphan crop example where the approach is beginning to be explored is finger millet (*Eleusine coracana*), a seed-propagated annual grain (discussed further below). As for standard marker-assisted selection, the primary limitation in applying genomic selection to new and orphan crops is the absence of phenotypic data from relevant training populations (Varshney *et al*., 2012). The application of the approach could in theory however be especially effective for slow-maturing perennial new and orphan crops which are difficult to directly phenotype for key production traits (Isik *et al*., 2015). Genomic selection may also be particularly effective when the underlying biological basis of key traits is poorly understood, as is the case with many new and orphan crops.

Along with the advent of new technological approaches for crop genetic improvement there is an emerging reconsideration of breeding methods that effectively address spatial and temporal interactions among the different biological components in diverse production systems (Litrico & Violle, 2015). The identification of target crop-crop interaction traits in this breeding is partially informed by research on natural systems, as outlined earlier in this review (Section III.2). In agricultural systems, however, it is also possible to 'force' positive relationships among crop diversity, yield and overall production stability that are unrealisable naturally due to different balancing trade-offs (Denison *et al*., 2003). Significant research is though still required to explore context-appropriate intercrop breeding techniques; the approaches to intercrop breeding that have been proposed so far have rarely been implemented (Hamblin *et al*., 1976; Wright, 1985).

Theoretically, genomic selection offers clear advantages for intercrop breeding as it can better manage the expected complex genetics of interaction traits and it reduces the need for large experimental plots to evaluate crop-crop interactions. In our own research (JB, JH, SH, IKD, in collaboration with the International Crops Research Institute for the Semi-Arid Tropics) we are stochastically modelling the effectiveness of intercrop breeding with and without the application of genomic selection for finger millet and the accompanying legume crop of groundnut (*Arachis hypogaea*) (Fig. 5). These crops are grown together in East Africa within low input smallholder production systems that can benefit significantly from exploiting crop-crop synergies (Yu *et al*., 2015). Better varietal combinations of finger millet and groundnut could support higher yielding, more stable and more sustainable agricultural production in the region. The importance of cereal-legume combinations globally means that our modelling also has broader application.

Speed breeding, which reduces the generation interval in breeding programmes by altering the photoperiod exposure of day length sensitive plants to accelerate their development (normally through prolonging "long-day" plants' exposure to light; Ghosh *et al*., 2018), is another approach now being applied to orphan crops. Application includes to the long-day annual legume chickpea (*Cicer arietinum*), for which the number of possible generations per year has been increased from three to six (Watson *et al*., 2018). Application has also been successful with "short-day" annual grain amaranth crops (*Amaranthus* spp.; Stetter *et al*., 2016). Speed breeding should be especially effective in combination with genomic selection as this allows selection during rapid cycling where full phenotypic data are not collected (Li *et al*., 2018a). It has been proposed that the costs of the speed breeding approach for new and orphan crops in low income nations could be reduced through transportable “speed breeding capsules”, consisting of shipping containers retrofitted with temperature and light controls, irrigation systems and greenhouse benches (Chiurugwi *et al*., 2019).

Participatory breeding and selection

“Citizen science” projects that evaluate crop germplasm have been conducted successfully in high income nations, as illustrated by Würschum *et al*. (2019) who explored genotype-environment interactions in soybean (*Glycine max*) based on data collected by 1,800 gardeners located across Germany. Even higher potential for participatory experimentation exists in low income nations where rural populations are greater and include many active small-scale farmers. If properly supported, these communities can meaningfully evaluate genetic materials within a range of target environments and cropping systems, and provide further information on crops’ production and consumption.

In Central Africa, for example, participatory domestication methods have been successfully applied to genetically improve new and orphan fruit tree crops including the semi-domesticated safou (*Dacryodes edulis*) and the incipiently domesticated bush mango (*Irvingia gabonensis* and *I. wombolu*) (Jamnadass *et al*., 2011). Here, scientific advances in tree selection, propagation and fruit processing were combined with local communities’ experiences in tree management. Applying simple selection methods and basic vegetative propagation approaches resulted in significant yield and quality gains from existing wide gene pools of these fruit trees and the effective fixation of these polygenic traits (Tchoundjeu *et al*., 2006). Vegetative propagation also significantly reduced the interval between crop establishment and production, and produced smaller, easier to harvest, plants, thereby increasing returns to farmers’ labour. By linking production to processing and market development, the participatory tree domestication approach has spread in the Central Africa region (Asaah *et al*., 2011).

Participatory approaches have also been applied in the Middle East and North Africa to cereals, combining centralised and decentralised breeding through the deployment to farmers of crop germplasm panels assembled by breeders and scientists (Mustafa *et al*., 2006). The participatory approach is considered especially useful for the heterogeneous production conditions common in low income nations and when the preference for specific crop traits is poorly understood (Bhargav & Meena, 2014). Both of these conditions often apply for new and orphan crops.

Environment-based selection

“Landscape genomic” approaches to crop development are particularly relevant for perennial plants that exist currently mostly as wild populations adapted over many generations to local abiotic conditions (Bragg *et al*., 2015). This is because the ‘*in situ*’ decision making that is involved avoids the considerable time and effort required to evaluate perennial crop germplasm in formal field trials. In the approach, genomic data collected from plants growing in natural populations are correlated with environmental variables using statistical methods that account for underlying adaptively neutral genetic structure caused by genetic drift (Coop *et al*., 2010). Established correlations can then, in theory, be used to screen wider germplasm panels to determine favourable allele compositions for particular production conditions. Comparisons are facilitated by the large number of georeferenced interpolated environmental data sets now available digitally, including temperature and precipitation profiles (e.g., Fick & Hijmans, 2017) and soil types (e.g., ISRIC, 2019). In a landscape genomic analysis covering the native range of barrel medic (*Medicago truncatula*, a legume), for example, Guerrero *et al*. (2018) made use of soil maps to identify soil environment as a key driver of adaption, with a high number of SNPs associated with soil variables, including SNPs in candidate genes involved in nodulation/symbiotic nitrogen fixation.

If local adaptation can be assumed to have occurred during orphan crop development and ecogeographic range expansion, the landscape approach can be applied to orphan crop landraces as well as to new and orphan crops’ wild germplasm. In this case, meta-analysis of multiple crops’ progenitors and landraces in the same geographic space could provide comparative insights into mechanisms of natural and human adaptation. Statistical approaches are now available that combine the results of multi-common garden genome-wide association studies, which explore the genetic basis of phenotype-trial site interactions, with wild and/or landrace sample environment-genomic correlations (Lasky *et al*., 2018). This can further facilitate an understanding of causal loci for adaptation and help to define appropriate strategies for new and orphan crops’ range expansions.

IV.2. Orthologous gene involvement in new and orphan crop trait evolution

An understanding of the extent to which the evolution of a common phenotype among existing crops has involved mutations in orthologous gene sequences as opposed to changes in different genes is of clear practical relevance for new and orphan crop development (Pickersgill, 2018). Clearly, the greater the extent of orthologous gene involvement in common trait evolution in past crop domestications, then the more attractive it is to target change to related gene sequences in new and orphan crops to drive their domestication forward. Indeed, the important roles of orthologous gene sequences in crops' domestications have been widely revealed (Martin & Orgogozo, 2013). For example, orthologous sequences control at least a portion of variation in flowering time (Calixto *et al*., 2015), plant height (Jia *et al*., 2009), grain stickiness (Meyer & Purugganan, 2013), seed size (Tao *et al*., 2017), seed dormancy (Wang *et al*., 2018) and seed and fruit dispersal or retention (Li & Olsen, 2016) across various crops (see examples in Table 1). As variation for a number of these traits underlies new and orphan crop development priorities, focusing on relevant gene orthologues, defined by comparisons with suitable crop exemplars (further addressed in Section IV.3), is clearly of value.

As would be expected, in general the more closely related two crops are then the more likely they are to share the same underlying genes and genetic architectures for in-common phenotypes (Lenser & Theißen, 2013). However, orthologous genes are involved in determining common phenotypes even when crops are evolutionarily distant, as is evident from some of the examples in Table 1 of our current review. For example, allelic variants in orthologues of the rice *WAXY* gene control starch composition not only in a range of cereals and pseudo-cereals from Poaceae to Amaranthaceae (see Box 2 in Meyer & Purugganan, 2013), but even in some non-cereal crops (e.g., Wang *et al*., 2017). On the other hand, multiple domestications within a single crop species may involve unrelated genes to reach a common phenotype (Meyer & Purugganan, 2013), illustrating the breadth of possible mechanisms involved in crop evolution and that assumptions of orthology should be guarded.

In general, the literature suggests that the genes associated with initial domestication processes are more in common across crops than the genes associated with diversification (Lai *et al*., 2018; though see discussion in Pickersgill, 2018). This would suggest that the exploitation of crop-crop orthologous gene relationships could be more beneficial when domesticating entirely new crops (*de novo* domesticates) rather than when further developing orphan crops that have already passed through initial domestication barriers. Importantly, however, even if in many cases crops' common phenotypes were reached via alterations in unrelated genes in past domestications, this does not preclude the targeting of orthologous sequences in the further domestication of orphan crops, as a transgressive approach from previous domestication pathways may still prove effective and could be more efficient (Lenser & Theißen, 2013). Indeed, the use of advanced molecular breeding methods such as gene editing to effect changes in domestication-related gene orthologues has been shown to be effective for orphan crops in some circumstances: for example, recent research on the solanaceous orphan crop groundcherry (*Physalis pruinosa*) using CRISPR/Cas9 to mutate orthologues of tomato domestication and improvement genes has shown promise (Lemmon *et al*., 2018).

Clearly, the effectiveness of different breeding approaches will depend on the varying underlying basis of traits' evolution (Østerberg *et al*., 2017). Applying knockout CRISPR/Cas9 gene editing to new and orphan crop gene orthologues of known cross-crop, large effect, initial 'domestication' genes, for which change has often been associated with loss-of-function mutations, seems advisable. In contrast, application to smaller effect 'optimisation' or 'diversification' genes, where a less clear orthology exists and where change has more often been associated with gains in function, seems less advisable. The reduction or removal of anti-nutrients via CRISPR/Cas9 or other mutational (e.g., TILLING) disruption of dedicated orthologous genes in conserved metabolic pathways (e.g., Emmrich, 2017) could also be particularly effective.

IV.3. Identifying exemplar crops to inform new and orphan crop domestication pathways

To determine appropriate genetic improvement pathways for new and orphan crops the development routes of more widely researched crops should be considered. If orthologous approaches to improvement are to be best exploited, identifying a given new or orphan crop's most relevant more widely studied exemplar requires considering the genetic relatedness of crop pairs. The definition of exemplars also requires consideration of crops' biologies. To illustrate how appropriate exemplars may be identified on the basis of genetic relatedness and crop biology, we have compared a group of exemplar-requiring new and orphan crops with a panel of possible crop models using taxonomy as a proxy for genetic relatedness (as described in Supporting Information Notes S3; raw data for analysis provided in Supporting Information Table S1). In our analysis, possible exemplars include major crops and new and orphan crops chosen based on the availability of crop production, trade and food balance data in FAOSTAT (2019) databases. These data provide some contextual understanding of recent crop development (e.g., see Fig. 3). The exemplars chosen also represent a range of production biologies, including perenniality and vegetative propagation. In our comparison, the exemplar-requiring new and orphan crops we chose are all considered important for supporting human nutrition in Africa. The results of our analysis, illustrated in Figure 6 (detailed results provided in Supporting Information Table S1), revealed promising pairings between exemplar-requiring and exemplar crops that were not always intuitive. In addition, depending on the relative emphasis given to biology and taxonomy in the analysis (adjusting from a 1:1 biology:taxonomy weighting to 2:1 or 1:2 weightings when calculating paired crop distances), some change in pairings was evident.

Greater insights into new and orphan crop genomes are emerging from current sequencing efforts. As only one example, the African Orphan Crops Consortium is assembling genomes and resequencing representative germplasm panels for 101 new and orphan crops (AOCC, 2019). These plants represent a prioritised list of exotic and indigenous species to Africa that are important for meeting human nutritional needs and providing other services that support farmers' livelihoods on the continent. These other services include those that are not directly provisioning, such as environmental services, where crop interactions with each other and with other biotic components of farm landscapes are important. With such sequencing efforts underway, it is possible to focus more intently on questions of genetic relatedness in efforts to identify new exemplar crops that have the potential to support transgressive, orthologue-based approaches to domestication.

V. Future outlook

Comparative research on crops along a domestication continuum allows the value of different genetic improvement approaches to be determined (Fig. 1). Should more focus, for example, be placed on environmentally-based *in situ* selection strategies for wild relatives and progenitors of major crops? And how much more effort should be given to develop *ex situ* collections of potential new crops, in order to apply advanced and/or conventional breeding approaches to them? In particular, it is evident that the central bridging positioning of orphan crops provides unique opportunities for investigating genetic improvement approaches that both support *de novo* domestications and major crop 'rewildings' (*sensu* Palmgren *et al*., 2015). In the case of new domestications, for example, it is important to understand how effectively crop development can be driven by the knockout of candidate domestication-related genes, using modern gene editing technologies (Østerberg *et al*., 2017). If a comparison of the gene sequences of an orphan crop's widely prevalent extant wild progenitors and farmed semi-domesticates indicates that the latter's development was based on loss-of-function mutations of fundamental domestication genes, a knockout approach to domestication could be successfully applied to the wide extant wild germplasm base of (putative) new crops and could, if desired, be first further practically tested on orphan crops.

Further practical testing of the above approach would in addition reveal if there is merit in the 'redomestication' of major crops from their wild relatives and progenitors as a strategy for efficiently accessing wild gene pools for traits lost in the development of advanced cultivars but now considered beneficial for addressing agriculture's sustainability challenges (Langridge & Waugh, 2019). Recent research using CRISPR/Cas9 gene editing of target domestication-related genes has shown promise for redomestications, with domesticated phenotypes that retain important wild attributes achievable starting from crop wild progenitors in the case of tomato (Li *et al*., 2018b; Zsögön *et al*., 2018). It is known that wild relatives, progenitors and landraces of a number of major crops contain more variation in traits related to resource use efficiency and a plant's ability to interact positively with other crops and non-crop biotic components in complex production systems than do narrowly-diverse advanced cultivars developed for monoculture (Kapulnik & Kushnir, 1991; Mutch & Young, 2004; Martín-Robles *et al*., 2018). Rewilding major crops for these traits (Palmgren *et al*. 2015), sampling variation at relevant gene sequences whose identification is supported by new and orphan crop analysis (Jacob *et al*., 2018), could then be an effective approach for sustainably intensifying farming, especially when crop interactions are specifically considered in breeding (Litrico & Violle, 2015). At the same time, ensuring that these once-cryptic sustainability features are maintained in new and orphan crops as their domestication either begins or intensifies is clearly important for ensuring more holistic farming system improvement outcomes (Dawson *et al*., 2018a).

Acknowledgements

Our thanks to the many colleagues with whom we have had discussions on new and orphan crop development. SRUC authors of this review gratefully acknowledge Global Challenge Research Funding on orphan crops (project BB/P022537/1: Formulating Value Chains for Orphan Crops in Africa, 2017-2019, Foundation Award for Global Agriculture and Food Systems). World Agroforestry (ICRAF) authors gratefully acknowledge the support of the CGIAR's funding partners for their work (https://www.cgiar.org/funders/).

References

Alonso C, Ramos-Cruz D, Becker C. 2019. The role of plant epigenetics in biotic interactions. *New Phytologist* **221:** 731-737.

AOCC. 2019. *The African Orphan Crops Consortium*. [WWW site] URL http://africanorphancrops.org/ [accessed 10 January 2019].

Asaah EK, Tchoundjeu Z, Leakey RRB, Takousting B, Njong J, Edang I. 2011. Trees, agroforestry and multifunctional agriculture in Cameroon. *International Journal of Agricultural Sustainability* **9:** 110-119.

Asano K, Takashi T, Miura K, Qian Q, Kitano H, Matsuoka M, Ashikari M. 2007. Genetic and molecular analysis of utility of *sd1* alleles in rice breeding. *Breeding Science* **57:** 53-58.

Azzi L, Deluche C, Gévaudant F, Frangne N, Delmas F, Hernould M, Chevalier C. 2015. Fruit growth-related genes in tomato. *Journal of Experimental Botany* **66:** 1075-1086.

Bakare AH, Osundahunsi OF, Olusanya JO. 2016. Rheological, baking, and sensory properties of composite bread dough with breadfruit (*Artocarpus communis* Forst) and wheat flours. *Food Science & Nutrition* **4:** 573-587.

Ballaré CL, Pierik R. 2017. The shade-avoidance syndrome: multiple signals and ecological consequences. *Plant, Cell & Environment* **40:** 2530-2543.

Balmer M. 2007. Energy poverty and cooking energy requirements: the forgotten issue in South African energy policy? *Journal of Energy in Southern Africa* **18:** 4-9.

Bergonzi S, Albani MC. 2011. Reproductive competence from an annual and a perennial perspective. *Journal of Experimental Botany* **62:** 4415-4422.

Bhargav DK, Meena HP. 2014. Participatory plant breeding: farmers as breeders. *Popular Kheti* **2:** 7-14.

Bragg JG, Supple MA, Andrew RL, Borevitz JO. 2015. Genomic variation across landscapes: insights and applications. *New Phytologist* **207:** 953-967.

Brooker RW, Bennett AE, Cong W-F, Daniell TJ, George TS, Hallett PD, Hawes C, Iannetta PPM, Jones HG, Karley AJ *et al*. 2015. Improving intercropping: a synthesis of research in agronomy, plant physiology and ecology. *New Phytologist* **206:** 107-117.

Bustos-Korts D, Dawson IK, Russell J, Tondelli A, Guerra D, Ferrandi C, Strozzi F, Nicolazzi EL, Molnar-Lang M, Ozkan H *et al*. 2019. Exome sequences and multi-environment field trials elucidate the genetic basis of adaptation in barley. *The Plant Journal* (in press).

Calixto, CPG, Waugh R, Brown JWS. 2015. Evolutionary relationships among barley and *Arabidopsis* core circadian clock and clock-associated genes. *Journal of Molecular Evolution* **80:** 108-119.

Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, Narwani A, Mace GM, Tilman D, Wardle DA *et al*. 2012. Biodiversity loss and its impact on humanity. *Nature* **486:** 59-67.

CFF. 2019. *Crops for the Future. Facilitating the wider use of underutilised crops.* [WWW site] URL http://www.cffresearch.org/ [accessed 16 April 2019].

Chakrabarti M, Zhang N, Sauvage C, Muños S, Blanca J, Cañizares J, Diez MJ, Schneider R, Mazourek M, McClead J *et al*. 2013. A cytochrome P450 regulates a domestication trait in cultivated tomato. *Proceedings of the National Academy of Sciences USA* **110:** 17125-17130.

Chiurugwi T, Kemp S, Powell W, Hickey LT. 2019. Speed breeding orphan crops. *Theoretical and Applied Genetics* **132:** 607-616.

Cichy KA, Wiesinger JA, Mendoza FA. 2015. Genetic diversity and genome‐wide association analysis of cooking time in dry bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics* **128:** 1555-1567.

Clay JW. 2004. *World agriculture and the environment: a commodity-by-commodity guide to impacts and practices*. Washington DC, USA: Island Press.

Clement CR. 2004. Fruits. In: Prance GT, Nesbitt M, eds. *The cultural history of plants*. London, UK: Routledge, 77-95.

Comadran J, Kilian B, Russell J, Ramsay L, Stein N, Ganal M, Shaw P, Bayer M, Thomas W, Marshall D *et al*. 2012. Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nature Genetics* **44:** 1388-1392.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK. 2010. Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185:** 1411-1423.

Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, de los Campos G, Burgueño J, González-Camacho JM, Pérez-Elizalde S, Beyene Y. 2017. Genomic selection in plant breeding: methods, models, and perspectives. *Trends in Plant Science* **22:** 961-975.

Darwin C. 1859. *On the origin of species by means of natural selection, or the preservation of favoured races in the* struggle for life. London, UK: John Murray.

Dawson IK, Attwood SJ, Park SE, Jamnadass R, Powell W, Sunderland T, Kindt R, McMullin S, Hoebe PN, Baddeley J *et al*. 2018a. *Contributions of biodiversity to the sustainable intensification of food production. Thematic study for the State of the World's Biodiversity for Food and Agriculture.* Rome, Italy: Food and Agriculture Organization of the United Nations.

Dawson IK, Hendre P, Powell W, Sila D, McMullin S, Simons T, Revoredo-Giha C, Odeny DA, Barnes AP, Graudal L *et al*. 2018b. *Supporting human nutrition in Africa through the integration of new and orphan crops into food systems: placing the work of the African Orphan Crops Consortium in context*. ICRAF Working Paper No. 276. Nairobi, Kenya: The World Agroforestry Centre.

Denison RF, Kiers ET, West SA. 2003. Darwinian agriculture: when can humans find solutions beyond the reach of natural selection? *The Quarterly Review of Biology* **78:** 145-168.

Doebley J, Stec A, Hubbard L. 1997. The evolution of apical dominance in maize. *Nature* **386:** 485-488.

Druet T, Macleod IM, Hayes BJ. 2014. Toward genomic prediction from whole-genome sequence data: impact of sequencing design on genotype imputation and accuracy of predictions. *Heredity* **112:** 39-47.

Economist. 2017. *No crop left behind: improving the plants that Africans eat and breeders neglect*. The Economist. [WWW document] URL http://www.economist.com/news/science-and-technology/21731614-nutrition-and-genetics-africa-improving-plants-africans-eat-and [accessed 5 December 2018].

Emmrich PMF. 2017. Genetic improvement of grass pea (*Lathyrus sativus*) for low β-L-ODAP content. PhD thesis. Norwich, UK: John Innes Centre.

FAOSTAT. 2019. *Food and Agriculture Organization of the United Nations Statistics Division portal*. [WWW site] URL http://faostat3.fao.org/ [accessed 15 January 2019].

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37:** 4302-4315.

Foster J, Kim HU, Nakata PA, Browse J. 2012. A previously unknown oxalyl-CoA synthetase is important for oxalate catabolism in *Arabidopsis*. *The Plant Cell* **24:** 1217-1229.

Francis CA, Smith ME. 1985. Variety development for multiple cropping systems. *Critical Reviews in Plant Sciences* **3:** 133-168.

Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD. 2000. *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* **289:** 85-88.

Gamuyao R, Chin JH, Pariasca-Tanaka J, Pesaresi P, Catausan S, Dalid C, Slamet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S. 2012. The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature* **488:** 535-539.

Georges M, Charlier C, Hayes B. 2019. Harnessing genomic information for livestock improvement. *Nature Reviews Genetics* **20:** 135-156.

Ghosh S, Watson A, Gonzalez-Navarro OE, Ramirez-Gonzalez RH, Yanes L, Mendoza-Suárez M, Simmonds J, Wells R, Rayner T, Green P *et al*. 2018. Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nature Protocols* **13:** 2944-2963.

Gillespie S, van den Bold M. 2017. Agriculture, food systems, and nutrition: meeting the Challenge. *Global Challenges* **1:** Article 1600002. URL https://onlinelibrary.wiley.com/doi/epdf/10.1002/gch2.201600002

Gruber K. 2017. The living library. *Nature* **544:** S8-S10.

Goldschmidt EE. 2013. The evolution of fruit tree productivity: a review. *Economic Botany* **67:** 51-62.

Guerrero J, Andrello M, Burgarella C, Manel S. 2018. Soil environment is a key driver of adaptation in *Medicago truncatula*: new insights from landscape genomics. *New Phytologist* **219:** 378-390.

Guseman JM, Webb K, Srinivasan C, Dardick C. 2017. *DRO1* influences root system architecture in Arabidopsis and Prunus species. *The Plant Journal* **89:** 1093-1105.

Halewood M, Chiurugwi T, Sackville Hamilton R, Kurtz B, Marden E, Welch E, Michiels F, Mozafari J, Sabran M, Patron N *et al*. 2018. Plant genetic resources for food and agriculture: opportunities and challenges emerging from the science and information technology revolution. *New Phytologist* **217:** 1407-1419.

Hamblin J, Rowell JG, Redden R. 1976. Selection for mixed cropping. *Euphytica* **25:** 97-106.

Hayes BJ, Bowman PJ, Chamberlain AC, Verbyla K, Goddard ME. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genetics Selection Evolution* **41:** Article 51. URL https://doi.org/10.1186/1297-9686-41-51

Hickey JM, Chiurugwi T, Mackay I, Powell W, Eggen A, Kilian A, Jones C, Canales C, Grattapaglia D, Bassi F *et al*. 2017. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature Genetics* **49:** 1297-1303.

Isik F, Kumar S, Martínez-García PJ, Iwata H, Yamamoto T. 2015. Acceleration of forest and fruit tree domestication by genomic selection. *Advances in Botanical Research* **74:** 93-124.

ISRIC. 2019. *ISRIC – world soil information*. [WWW site] URL https://www.isric.org/ [Lasted accessed 14 December 2018].

Iwata H, Minamikawa MF, Kajiye-Kanegae H, Ishimori M, Hayashi T. 2016. Genomics-assisted breeding in fruit trees. *Breeding Science* **66:** 100-115.

LATINCROP. 2019. *LATINCROP. An integrated strategy for the conservation and use of underutilized Latin American agrobiodiversity* [WWW site] URL http://www.latincrop.org/ [accessed 16 April 2019].

Lin Z, Li X, Shannon LM, Yeh C-T, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE *et al*. 2012. Parallel domestication of the *Shattering1* genes in cereals. *Nature Genetics* **44:** 720-724.

Jacob P, Avni A, Bendahmane A. 2018. Translational research: exploring and creating genetic diversity. *Trends in Plant Science* **23:** 42-52.

Jamnadass RH, Dawson IK, Franzel S, Leakey RRB, Mithöfer D, Akinnifesi FK, Tchoundjeu Z. 2011. Improving livelihoods and nutrition in sub-Saharan Africa through the promotion of indigenous and exotic fruit production in smallholders' agroforestry systems: a review. *International Forestry Review* **13:** 338-354.

Jia G, Huang X, Zhi H, Zhao Y, Zhao Q, Li W, Chai Y, Yang L, Liu K, Lu H *et al*. 2013.

A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nature Genetics* **45:** 957-961.

Jia Q, Zhang J, Westcott S, Zhang X-Q, Bellgard M, Lance R, Li C. 2009. GA-20 oxidase as a candidate for the semidwarf gene *sdw1*/*denso* in barley. *Functional & Integrative Genomics* **9:** 255-262.

Jones PMB, Boulter D. 1983. The analysis of development of hardbean during storage of black beans (*Phaseolus vulgaris L*). *Plant Foods for Human Nutrition* **33:** 77-85.

Kaur B, Ranawana V, Henry J. 2016. The glycaemic index of rice and rice products: a review, and table of GI values. *Critical Reviews in Food Science and Nutrition* **56:** 215-236.

Kessides C. 2005. *The urban transition in Sub-Saharan Africa: implications for economic growth and poverty reduction*. Africa Region, Working Paper Series No. 97. Washington DC, USA: The World Bank.

Khoury CK, Bjorkman AD, Dempewolf H, Ramirez-Villegas J, Guarino L, Jarvis A, Rieseburg LH, Struik PC. 2014. Increasing homogeneity in global food supplies and the implications for food security. *Proceedings of the National Academy of Sciences USA* **111:** 4001-4006.

Khoury CK, Jarvis A. 2014. *The changing composition of the global diet: implications for CGIAR research*. CIAT Policy Brief No. 18. Cali, Colombia: Centro Internacional de Agricultura Tropical.

Kotoda N, Iwanami H, Takahashi S, Abe K. 2006. Antisense expression of *MdTFL1*, a *TFL1*-like gene, reduces the juvenile phase in apple. *Journal of the American Society for Horticultural Science* **131:** 74-81.

Lagassé SL, Hatcher DW, Dexter JE, Rossnagel BG, Izydorczyk MS. 2006. Quality characteristics of fresh and dried white salted noodles enriched with flour from hull-less barley genotypes of diverse amylose content. *Cereal Chemistry* **83:** 202-210.

Lai X, Yan L, Lu Y, Schnable JC. 2018. Largely unlinked gene sets targeted by selection for domestication syndrome phenotypes in maize and sorghum. *Plant Journal* **93:** 843-855.

Langridge P, Waugh R, 2019. Harnessing the potential of germplasm collections. *Nature Genetics* **51:** 200-201.

Larson G, Piperno DR, Allaby RG, Purugganan MD, Andersson L, Arroyo-Kalin M, Barton L, Climer Vigueira C, Denham T, Dobney K *et al*. 2014. Current perspectives and the future of domestication studies. *Proceedings of the National Academy of Sciences USA* **111:** 6139-6146.

Lasky JR, Forester BR, Reimherr M. 2018. Coherent synthesis of genomic associations with phenotypes and home environments. *Molecular Ecology Resources* **18:** 91-106.

Lemmon ZH, Reem NT, Dalrymple J, Soyk S, Swartwood KE, Rodriguez-Leal D, Van Eck J, Lippman ZB. 2018. Rapid improvement of domestication traits in an orphan crop by genome editing. *Nature Plants* **4:** 766-770.

Lenser T, Theißen G. 2013. Molecular mechanisms involved in convergent crop domestication. *Trends in Plant Science* **18:** 704-714.

Li H, Rasheed A, Hickey LT, He Z. 2018a. Fast-forwarding genetic gain. *Trends in Plant Science* **23:** 184-186.

Li L-F, Olsen KM. 2016. To have and to hold: selection for seed and fruit retention during crop domestication. *Current Topics in Developmental Biology* **119:** 63-109.

Li T, Yang X, Yu Y, Si X, Zhai X, Zhang H, Dong W, Gao C, Xu C. 2018b. Domestication of wild tomato is accelerated by genome editing. *Nature Biotechnology* **36:** 1160-1163.

Litrico I, Violle C. 2015. Diversity in plant breeding: a new conceptual framework. *Trends in Plant Science* **20:** 604-613.

Liu H-J, Yan J. 2019. Crop genome-wide association study: a harvest of biological relevance. *The Plant Journal* **97**: 8-18.

MacLeod IM, Bowman PJ, Vander Jagt CJ, Haile-Mariam M, Kemper KE, Chamberlain AJ, Schrooten C, Hayes BJ, Goddard ME. 2016. Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* **17:** Article 144. URL https://doi.org/10.1186/s12864-016-2443-6

Mai CD, Phung NTP, To HTM, Gonin M, Hoang GT, Nguyen KL, Do VN, Courtois B, Gantel P. 2014. Genes controlling root development in rice. *Rice* **7:** Article 30. URL https://doi.org/10.1186/s12284-014-0030-5

Martin A, Orgogozo V. 2013. The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* **67:** 1235-1250.

Martín-Robles N, Lehmann A, Seco E, Aroca R, Rillig MC, Milla R. 2018. Impacts of domestication on the arbuscular mycorrhizal symbiosis of 27 crop species. *New Phytologist* **218:** 322-334.

Meuwissen TH, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157:** 1819-1829.

Meyer RS, DuVal AE, Jensen HR. 2012. Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytologist* **196:** 29-48.

Meyer RS, Purugganan MD. 2013. Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics* **14:** 840-852.

Migicovsky Z, Myles S. 2017. Exploiting wild relatives for genomics-assisted breeding of perennial crops. *Frontiers in Plant Science* **8:** Article 460. URL https://doi.org/10.3389/fpls.2017.00460

Miller AJ, Gross BL. 2011. From forest to field: perennial fruit crop domestication. *American Journal of Botany* **98:** 1389-1414.

Mustafa MA, Mayes S, Massawe F. 2019. Crop diversification through a wider use of underutilised crops: a strategy to ensure food and nutrition security in the face of climate change. In: Sarkar A, Sensarma SR, vanLoon GW, eds. *Sustainable solutions for food security: combating climate change by adaptation*. Springer Nature Switzerland AG, Cham, Switzerland, 125-149.

Mustafa Y, Grando S, Ceccarelli S. 2006. *Assessing the benefits and costs of participatory and conventional barley breeding programs in Syria*. Aleppo, Syria: The International Center for Agricultural Research in the Dry Areas.

Nakamichi N. 2015. Adaptation to the local environment by modifications of the photoperiod response in crops. *Plant Cell Physiology* **56:** 594-604.

Nour-Eldin HH, Andersen TG, Burow M, Madsen SR, Jørgensen ME, Olsen CE, Dreyer I, Hedrich R, Geiger D, Halkier BA. 2012. NRT/PTR transporters are essential for translocation of glucosinolates defence compounds to seeds. *Nature* **488:** 531-534.

Nour-Eldin HH, Halkier BA. 2013. The emerging field of transport engineering of plant specialized metabolites. *Current Opinion in Biotechnology* **24:** 263-270.

Palmgren MG, Edenbrandt AK, Vedel SE, Andersen MM, Landes X, Østerberg JT, Falhof J, Olsen LI, Christensen SB, Sandøe P *et al*. 2015. Are we ready for back-to-nature crop breeding? *Trends in Plant Science* **20:** 155-164.

Panzeri D, Cassani E, Doria E, Tagliabue G, Forti L, Campion B, Bollini R, Brearley CA, Pilu R, Nielsen E *et al*. 2011. A defective ABC transporter of the MRP family, responsible for the bean *lpa1* mutation, affects the regulation of the phytic acid pathway, reduces seed *myo*-inositol and alters ABA sensitivity. *New Phytologist* **191:** 70-83.

Petry N, Boy E, Wirth JP, Hurrell RF. 2015. The potential of the common bean (*Phaseolus vulgaris*) as a vehicle for iron biofortification, *Nutrients* **7:** 1144-1173.

Petry N, Egli I, Campion B, Nielsen E, Hurrell R. 2013. Genetic reduction of phytate in common bean (*Phaseolus vulgaris* l.) seeds increases iron absorption in young women. *The Journal of Nutrition* **143:** 1219-1224.

Pickersgill B. 2018. Parallel vs. convergent evolution in domestication and diversification of crops in the Americas. *Frontiers in Ecology and Evolution* **6:** Article 56. URL https://doi.org/10.3389/fevo.2018.00056

Popkin BM, Adair LS, Ng SW. 2012. Now and then: the global nutrition transition: the pandemic of obesity in developing countries. *Nutrition Reviews* **70:** 3-21.

Østerberg JT, Xiang W, Olsen LI, Edenbrandt AK, Vedel SE, Christiansen A, Landes X, Andersen MM, Pagh P, Sandøe P *et al*. 2017. Accelerating the domestication of new crops: feasibility and approaches. *Trends in Plant Science* **22:** 373-384.

Remigereau M-S, Lakis G, Rekima S, Leveugle M, Fontaine MC, Langin T, Sarr A, Robert T. 2011. Cereal domestication and evolution of branching: evidence for soft selection in the *Tb1* orthologue of pearl millet (*Pennisetum glaucum* [L.] R. Br.). *PLoS ONE* **6:** e22404. URL https://doi.org/10.1371/journal.pone.0022404

Russell J, Mascher M, Dawson IK, Kyriakidis S, Calixto C, Freund F, Bayer M, Milne I, Marshall-Griffiths T, Heinen S *et al*. 2016. Adaptation of barley to different environments revealed in the exomes of a range-wide collection of landraces and wild relatives. *Nature Genetics* **48:** 1024-1030.

Schilling S, Pan S, Kennedy A, Melzer R. 2018. MADS-box genes and crop domestication: the jack of all traits. *Journal of Experimental Botany* **69:** 1447-1469.

Schippmann U, Leaman DJ, Cunningham AB. 2002. The impact of cultivation and gathering of medicinal plants on biodiversity: global trends and issues. In: *Biodiversity and the ecosystem approach in agriculture, forestry and fisheries. Satellite event on the occasion of the Ninth Regular Session of the Commission on Genetic Resources for Food and Agriculture*. Rome, 12-13 October 2002. Rome, Italy: Food and Agriculture Organization of the United Nations, Inter-Departmental Working Group on Biological Diversity for Food and Agriculture.

Sheehan MJ, Kennedy LM, Costich DE, Brutnell TP. 2007. Subfunctionalization of *PhyB1* and *PhyB2* in the control of seedling and mature plant traits in maize. *The Plant Journal* **49:** 338-353.

Sousa DO, Carvalho AF, Oliveira JT, Farias DF, Castelar I, Oliveira HP, Vasconcelos IM. 2015. Increased levels of antinutritional and/or defense proteins reduced the protein quality of a disease-resistant soybean cultivar. *Nutrients* **7:** 6038-6054.

Stetter MG, Zeitler L, Steinhaus A, Kroener K, Biljecki M, Schmid KJ. 2016. Crossing methods and cultivation conditions for rapid production of segregating populations in three grain amaranth species. *Frontiers in Plant Science* **7:** Article 816. URL https://doi.org/10.3389/fpls.2016.00816

Studer AJ, Wang H, Doebley JF. 2017. Selection during maize domestication targeted a gene network controlling plant and inflorescence architecture. *Genetics* **207:** 755-765.

Subrahmaniam HJ, Libourel C, Journet E-P, Morel J-B, Munos S, Niebel A, Raffaele S, Roux F. 2018. The genetics underlying natural variation of plant-plant interactions, a beloved but forgotten member of the family of biotic interactions. *The Plant Journal* **93:** 747-770.

Sun-Waterhouse D, Zhao M, Waterhouse GIN. 2014. Protein modification during ingredient preparation and food processing: approaches to improve food processability and nutrition. *Food and Bioprocess Technology* **7:** 1853-1893.

Tadele Z. 2017. Raising crop productivity in Africa through intensification. *Agronomy* **7:** Article 22. URL https://doi.org/10.3390/agronomy7010022

Taketa S, Amano S, Tsujino Y, Sato T, Saisho D, Kakeda K, Nomura M, Suzuki T, Matsumoto T, Sato K *et al*. 2008. Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. *Proceedings of the National Academy of Sciences USA* **105:** 4062-4067.

Tao Y, Mace ES, Tai S, Cruickshank A, Campbell BC, Zhao X, Van Oosterom EJ, Godwin ID, Botella JR, Jordan DR. 2017. Whole-genome analysis of candidate genes associated with seed size and weight in *Sorghum bicolor* reveals signatures of artificial selection and insights into parallel domestication in cereal crops. *Frontiers in Plant Science* **8:** Article 1237. URL https://doi.org/10.3389/fpls.2017.01237

Tchoundjeu Z, Asaah E, Anegbeh PO, Degrande A, Mbile P, Facheux C, Tsobeng A, Atangana AR, Ngo-Mpeck ML, Simons AJ. 2006. Putting participatory domestication into practice in West and Central Africa. *Forests, Trees and Livelihoods* **16:** 53-69.

Turner A, Beales J, Faure S, Dunford RP, Laurie DA. 2005. The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* **310:** 1031-1034.

Uga Y, Sugimoto K, Ogawa S, Rane J, Ishitani M, Hara N, Kitomi Y, Inukai Y, Ono K, Kanno N *et al*. 2013. Control of root system architecture by *DEEPER ROOTING 1* increases rice yield under drought conditions. *Nature Genetics* **45:** 1097-1102.

UN. 2019. *United Nations Sustainable Development Goals*. United Nations. [WWW site] URL https://sustainabledevelopment.un.org/ [accessed 15 January 2019].

van der Knaap E, Chakrabarti M, Chu YH, Clevenger JP, Illa-Berenguer E, Huang Z, Keyhaninejad N, Mu Q, Sun L, Wang Y *et al*. 2014. What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. *Frontiers in Plant Science* **5:** Article 227. URL https://doi.org/10.3389/fpls.2014.00227

Vandermeer J. 1992. The ecology of intercropping. Cambridge, UK: Cambridge University Press.

Varshney RK, Ribaut J-M, Buckler ES, Tuberosa R, Rafalski JA, Langridge P. 2012. Can genomics boost productivity of orphan crops? *Nature Biotechnology* **30:** 1172-1176.

Varshney RK, Saxena RK, Upadhyaya HD, Khan AW, Yu Y, Kim C, Rathore A, Kim D, Kim J, An S *et al*. 2017. Whole-genome resequencing of pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits. *Nature Genetics* **49:** 1082-1088.

von Grebmer K, Saltzman A, Birol E, Wiesmann D, Prasai N, Yin S, Yohannes Y, Menon P, Thompson J, Sonntag A. 2014. *2014 Global hunger index: the challenge of hidden hunger*. Bonn, Germany: Deutsche Welthungerhilfe, with Washington DC, USA: International Food Policy Research Institute; and Dublin, Ireland: Concern Worldwide.

Walker T, Alene A, Ndjeunga J, Labarta R, Yigezu Y, Diagne A, Andrade R, Muthoni Andriatsitohaina R, De Groote H, Mausch K *et al*. 2014. *Measuring the effectiveness of crop improvement research in sub-Saharan Africa from the perspectives of varietal output, adoption, and change: 20 crops, 30 countries, and 1150 cultivars in farmers' fields*. Report of the Standing Panel on Impact Assessment. Rome, Italy: CGIAR Independent Science and Partnership Council Secretariat.

Wang M, Li W, Fang C, Xu F, Liu Y, Wang Z, Yang R, Zhang M, Liu S, Lu S *et al*. 2018. Parallel selection on a dormancy gene during domestication of crops from multiple families. *Nature Genetics* **50:** 1435-1441.

Wang Y, Li Y, Zhang H, Zhai H, Liu Q, He S. 2017. A soluble starch synthase I gene, *IbSSI*, alters the content, composition, granule size and structure of starch in transgenic sweet potato. *Nature Scientific Reports* **7:** Article 2315. URL www.nature.com/articles/s41598-017-02481-x

Wang Z-Y, Zheng F-Q, Shen G-Z, Gao J-P, Snustad DP, Li M-G, Zhang J-L, Hong M-M. 1995. The amylose content in rice endosperm is related to the post-transcriptional regulation of the *waxy* gene. *The Plant Journal* **7:** 613-622.

Watson A, Ghosh S, Williams MJ, Cuddy WS, Simmonds J, Rey M-D, Md Hatta MA, Hinchliffe A, Steed A, Reynolds D *et al*. 2018. Speed breeding is a powerful tool to accelerate crop research and breeding. *Nature Plants* **4:** 23-29.

Weltzien E, Christinck A. 2017. Participatory breeding: developing improved and relevant crop varieties with farmers. In: Snapp S, Pound B, eds. *Agricultural systems: agroecology and rural innovation for development*. 2nd edition. Cambridge, Massachusetts, USA: Academic Press, 259-301.

Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D, DeClerck F, Wood A *et al*. 2019. Food in the Anthropocene: the EAT–*Lancet* Commission on healthy diets from sustainable food systems. *The Lancet* **393:** 447-492.

Wolfe MD, Del Carpio DP, Alabi O, Ezenwaka LC, Ikeogu UN, Kayondo IS, Lozano R, Okeke UG, Ozimati AA, Williams E *et al*. 2017. Prospects for genomic selection in cassava breeding. *Plant Genome* **10:** Article 0015. URL https://dl.sciencesocieties.org/publications/tpg/abstracts/10/3/plantgenome2017.03.0015

Wright AJ. 1985. Selection for improved yield in inter-specific mixtures or intercrops. *Theoretical and Applied Genetics* **69:** 399-407.

Würschum T, Leiser WL, Jähne F, Bachteler K, Miersch M, Hahn V. 2019. The soybean experiment '1000 Gardens': a case study of citizen science for research, education, and beyond. *Theoretical and Applied Genetics* **132:** 617-626.

Yamagishi N, Kishigami R, Yoshikawa N. 2014. Reduced generation time of apple seedlings to within a year by means of a plant virus vector: a new plant-breeding technique with no transmission of genetic modification to the next generation. *Plant Biotechnology Journal* **12:** 60-68.

Yerra S, Putta S, Kilari EK. 2015. Detoxification of ODAP in *Lathyrus sativus* by various food processing techniques. *Pharmaceutical and Biological Evaluations* **2:** 152-159.

Yu Y, Stomph T-J, Makowski D, van der Werf W. 2015. Temporal niche differentiation increases the land equivalent ratio of annual intercrops: a meta-analysis. *Field Crops Research* **184:** 133-144.

Zsögön A, Čermák T, Naves ER, Notini MM, Edel KH, Weinl S, Freschi L, Voytas DF, Kudla J, Peres LEP. 2018. *De novo* domestication of wild tomato using genome editing. *Nature Biotechnology* **36:** 1211-1216.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Supporting Information Notes S1. Analysing production contributors to changes in global crop output.

Supporting Information Notes S2. Surveying plant breeders to identify new and orphan crop production constraints.

Supporting Information Notes S3. Comparing exemplar-requiring new and orphan crops with model crops.

Supporting Information Table S1. Supporting data for 30 new and orphan crops requiring exemplar models and 30 exemplar crops.

Figure legends

Figure 1. Schematic of the distribution of genetic diversity for new, orphan and major crops, with related improvement method options. The distribution of genetic resources (triangles and rectangle) varies by the category of plant, with implications for the application of different genetic improvement methods. For example, whereas major crops are well represented in gene banks globally, new crops are not; but in their case significant genetic variation is often still extant in the wild, though sometimes this variation is threatened (Dawson *et al.*, 2018b). Orphan crops occupy an intermediate position in the distribution of genetic resources across location categories and in their position on the domestication continuum. This positioning provides unique opportunities for orphan crops in investigating the extrapolation domains of a range of crop genetic improvement approaches, for *de novo* domestications and major crop ‘rewildings’ (rewilding *sensu* Palmgren *et al.*, 2015: the reestablishment of beneficial wild type properties in crops).

Figure 2. The diversity of farming systems in which winner and loser crops in the global food system are produced, based on data for 20 crops. The relationship between the diversity (summarised as intercrop or retained natural diversity; *y*-axis) of typical production systems and the relative change in food importance over the last half century (*x*-axis) for crops is shown. Crops were assigned numeric scores for production system diversity (ranging between 0 and 2, where 0 = lowest diversity, typically monoculture production) and change in food importance (positive scores = more important, negative scores = less important) by Dawson *et al*. (2018b), where further information on method can be found. Briefly, in the case of food importance, scores were based on the longitudinal trend analysis of Khoury *et al*. (2014) of FAOSTAT annual global food supply balance sheets, with crops showing a wide range of changes in relative food importance over the last half century being chosen as representative samples. Point size represents current global production area, based on a 2009-2013 mean (for reference purposes, the actual value for wheat, the crop with the largest production area, is 220 million ha). A linear regression indicates a trend toward lower diversity systems for increasingly important crops.

Figure 3. The relationship between production contributors (yield and total area) and changes in global output for 35 crops for the period 1961 to 2013. Our analysis is described in Supporting Information Notes S1. Briefly, values on the *y*-axis are the slope coefficients of linear regressions of

yield (production per unit area) and total production area contributions to crop output over the annual time series 1961 to 2013 for each of the 35 crops. Values of > and < 0 on the *y*-axis indicate relatively greater contributions from yield than total area and vice versa to global output over the time series, respectively. Values on the *x*-axis are changes in total global output over the 1961 to 2013 period. Point size represents current annual global gross production value, based on a 2009-2013 mean (for reference purposes, the actual value for paddy rice, the crop with the largest value, is 191 billion USD [in constant 2004-2006 USD]). As expected, a linear regression indicates an overall negative relationship between increases in output and the proportional contribution of yield to output for our crop panel, showing that in general crop yields were unable to keep apace with output increases over the last half century, in particular when output increases were very large. A group of nine crops where yield contributions to changes in output appear markedly low (well below the trend line) are encircled (dashed red line). With the exception of apple, these crops are characterised by relatively low production values (< USD 10 billion, compared to the mean for the total crop panel of USD 25 billion). Eight of the encircled crops are also perennial. In addition, six are (generally) propagated vegetatively under cultivation (cloves, coconut and sunflower being the seed-propagated exceptions). In a global context of limited land availability and increasing sustainability needs, moving such below-trend crops closer to the trend line may be an important measure for diversifying crop production.

Figure 4. Results of a survey of new and orphan crop production constraints, based on responses given by 53 African plant breeders on 30 specific plants of nutritional importance in Sub-Saharan Africa. The survey is described in Supporting Information Notes S2. (a) Priority production constraints, classified as genetic/management or input constraints. Crop pest or disease attack followed by storage problems were the most mentioned high priority genetic and/or management constraints and lack of access to suitable planting material followed by lack of crop-specific knowledge the most mentioned priority input constraints. (b) Key traits for genetic improvement of new and orphan crops. Pest and disease tolerance or resistance followed by yield (*per se*, i.e., independent of other production factors affecting yield, such as pest and disease attack) were the traits most mentioned as priorities for improvement (the apparent discrepancy between pie charts [a] and [b], where 'yield' as a unique feature is identified less often in the former case, appears to reflect yield in pie chart [a] being subsumed into [improved] 'planting material'). Improvement in harvestability was the fifth most mentioned category. (c) Key agronomic management interventions for new and orphan crop production. Improvements in planting and/or establishment methods

followed by soil fertilisation measures were the most mentioned priority interventions. Seasonal timing of field activities (such as planting, weeding and harvesting) was the fourth most mentioned category for intervention, with the fifth being interventions to diversify production systems. (d) Variation in priorities by plant primary food product for specific constraints and interventions identified by breeders (letters in parentheses as identified in pie charts [a] to [c]). Values are shown as proportions of all responses, by food product category (F = fruit, L = leaf, R = root, S = seed; for further information on these findings, see Supporting Information Notes S2). (e) Breeders' views of the potential for successful intervention in genetic improvement and in adopting new management practices. Here, breeders were asked to rate the potential for each of the key traits for genetic improvement or priority management interventions they had identified in (b) and (c), respectively, which were given equal weight as categories in analysis.

Figure 5. Intercrop breeding for finger millet and groundnut improvement. A proposed design with three example cycling/selection methods currently being explored via stochastic modelling is given: Base = non-genomic selection breeding approach with recurrent selection of parents based on their phenotypes at general intercropping ability (GIA) 1 and GIA 2 stages; GSPYT = genomic selection applied at the monoculture preliminary yield trial (PYT) stage to select new parents; and GSDH = genomic selection applied at the doubled haploid (DH, homozygous plant) stage to select new parents. Both of the shown example genomic selection scenarios currently being tested include advancement of individuals based on their genomic estimated breeding values (GEBV) in PYT and GIA 1 stages; in the GSDH scenario, this is additionally done in the DH stage. At the GIA 2 stage, only the individuals of the respective species with the best overall combining ability are advanced to specific intercropping ability (SIA) stage 1. Probe = an outstanding genotype of the alternate species used in combined test plots to evaluate intercropping ability.

Figure 6. Nearest exemplar crops, based on biologies and taxonomies, for 30 new and orphan crops in need of breeding method models. Crops chosen as exemplars are shown on the left of the figure and model-requiring new and orphan crops on the right. Connecting lines between crop pairs signify the minimum (Gower) distances between each model-requiring new or orphan crop and exemplar crops (analysis described in Supporting Information Notes S3; raw data and detailed results provided in Supporting Information Table S1). If analysis revealed more than one exemplar crop equally close to a model-requiring new or orphan crop, then multiple pairings are shown. To ease visualisation,

the 30 exemplar-requiring new and orphan crops we chose are divided into three groups of ten crops, with different coloured connector lines indicating minimum distances between crop pairs for each group. Solid connector lines represent an initial 1:1 biology:taxonomy weighting in the distance analysis. If there were differences in crop pairings when 2:1 or 1:2 biology:taxonomy weightings were subsequently applied, these are indicated by dashed connectors. Thirty exemplar crops, five of which were new or orphan crops and 25 of which were other crops, were chosen as the panel of exemplars because of the availability of production data for these crops in FAOSTAT. These exemplars are drawn from the crops (or crop groups) chosen for production trend analysis in Figure 3. Additional exemplars not specifically named in Figure 3 represent cases in which data were grouped for crops in the earlier figure (pooled reporting), but where component crops could be treated separately in current crop-crop comparisons.

Table 1. Illustrative genes for important new- and orphan crop-specific trait categories that may be targets for crop development

Trait category	Examples of relevant genes/pathways
Production integration	
Plant architecture	Major genes determining plant height are some of the best studied in the crop literature. In barley, for example, mutations in the *SEMI-DWARF1 (sdw1)* gene encoding the enzyme gibberellin 20-oxidase 3, which is involved in gibberellin biosynthesis, reduce plant stature (Jia *et al*., 2009). Mutations in the orthologous gene in rice, *SD1*, have been crucial in modern semi-dwarf rice variety development, one of the most important crop breeding interventions associated with the Green Revolution (Asano *et al*., 2007).
	Several genes that regulate plant branching architecture have been identified, including *TEOSINTE BRANCHED1* (*TB1*), which belongs to the TCP family of transcriptional regulators, in maize (*Zea mays*) (Studer *et al*., 2017). Orthologues include *Pgtb1* in pearl millet (*Pennisetum glaucum*) (Remigereau *et al*., 2011). The expression of *TB1* in maize is higher than in its progenitor (teosinte), conferring reduced branching (Doebley *et al*., 1997).
	Genes determining root architecture in rice include *DEEPER ROOTING 1* (*DRO1*) and *PHOSPHORUS-STARVATION TOLERANCE 1* (*PSTOL1*) (Mai *et al*., 2014). *DRO1*, a member of the IGT gene family, effects the root gravitropic response, via a modulation of epidermal cell elongation. It increases the angle between roots and the horizontal, inducing deeper rooting. The introduction of *DRO1* into a shallow-rooting rice cultivar enabled the resulting line to avoid drought (Uga *et al*., 2013). Orthologues appear to control root development in a range of other plants (Guseman *et al*., 2017). The *PSTOL1* gene, which encodes a receptor-like cytoplasmic kinase, is absent from modern rice varieties. Inserted into modern lines, it enhances early root growth, conferring greater root length and root surface area, and contributing to increased phosphorous uptake (Gamuyao *et al*., 2012)

Seasonal phenology	Gene networks controlling flowering are well researched, especially in cereals. In barley, for example, variation at the *PHOTOPERIOD-H1* (*HvPPD-H1*) gene, which encodes a pseudo-response regulator, and at the earliness *per se* gene *CENTRORADIALIS* (*HvCEN*), which encodes a phosphatidylethanolamine-binding protein, controls the days to heading trait (Russell *et al*., 2016). Causal variation at both these genes has been explored (Turner *et al*., 2005 and Comadran *et al*., 2012, respectively) and the magnitude of the effect of different haplotypes has been determined across multiple environments, allowing genotype-environment interactions to be characterised (Bustos-Korts *et al*., 2019)
Light competition	*PHY* genes encoding phytochrome photoreceptors and involved in plant growth regulator biosynthesis are involved in response to plant competition that changes the red to far red light ratio (Ballaré & Pierik, 2017). In maize, *PHYB1* and *PHYB2* genes encode phytochromes of the PHYB family that contribute differently to the shade avoidance response (Sheehan *et al*., 2007)
Product processability	
Anti-nutritional compounds	Biosynthetic and degradation pathways are known in model plants and have been studied in some orphan crops (especially legumes). Changes in single genes are able to influence both absolute level and organ allocation within the plant (Nour-Eldin & Halkier, 2013). Targeting the genes of specialised transport proteins essential for the transport of secondary metabolites, such as orthologues of *ARABIDOPSIS THALIANA GLUCOSINOLATE TRANSPORTER-1* (*GTR1*) and *GTR2* that are essential for the transport of glucosinolate defence compounds, could eliminate anti-nutrients from edible plant parts (Nour-Eldin *et al*., 2012). In common bean, an ethyl methanesulphonate mutant with significantly lowered phytic acid levels in seeds is affected in an MRP type ABC transporter gene, *Pvmrp1*, that is required for phytic acid accumulation and is orthologous to arabidopsis (*Arabidopsis thaliana*) *AtMRP5/AtABCC5* and maize *ZmMRP4* (Panzeri *et al*., 2011). In grass pea (*Lathyrus sativus*), the biosynthetic pathway of the neurotoxin β-N-ozalyl-L-α, β-diaminopropanoic acid (ODAP), which is a

	structural analogue of endogenous glutamate neurotransmitters, is not fully understood. But candidate genes for targeting, including a gene similar to that coding for an oxalyl-CoA synthetase in arabidopsis named *ACYL-ACTIVATING ENZYME3* (*AtAAE3*) that could catalyse the penultimate reaction step in the biosynthesis of ODAP (Foster *et al*., 2012), are currently under evaluation (Emmrich, 2017)
The "cooking time trait"	Genome-wide association scans have identified SNPs associated with cooking time on a number of common bean chromosomes (Pv02, Pv03, and Pv06). Proximate sequences of interest on Pv06 included two similar to arabidopsis *Cation/H(+) Antiporter 3* **(***AtCHX3*) and *AtCHX4* that transport calcium, a mineral known to influence cooking time for dry beans (Cichy *et al.*, 2015)
Processability traits for food formulation	Variation in the amylose to amylopectin ratio in cereal starches that affects consumer preference-, digestion- and processing-related traits has been identified with mutations at the rice *WAXY* gene *GRANULE BOUND STARCH SYNTHASE I* (*OsGBSS1*) and at orthologous sequences in a range of grains (Meyer & Purugganan, 2013). Mutations at *WAXY* that affect transcript processing and reduce GBSS activity confer the sticky (waxy) rice phenotype (low amylose to amylopectin ratio) (Wang *et al*., 1995).
	Ease of hull removal is an important physical property of grain that can influence its processability (e.g., ability to mill). In barley, the free-threshing (naked) phenotype is controlled by the *Nud* gene on chromosome 7H that encodes an ethylene response factor (ERF) family transcription factor involved in lipid biosynthesis. Deletion or low expression of the *Nud* gene results in the naked phenotype (Taketa *et al*., 2008)
Labour costs production	
Seed/fruit retention	The loss of seed and fruit dispersal mechanisms, which greatly facilitates harvesting efficiency, are key domestication syndrome traits (Meyer & Purugganan, 2013). Orthology is observed for some genes across crops, such as for *Shattering1* (*Sh1*), which encodes a YABBY transcription factor that provides shattering resistance in maize, sorghum (*Sorghum bicolor*) and rice. In domesticated sorghum, for

	example, a range of different types of mutations in *SbSh1* have led to reduced gene function and a reduction in shattering (Lin *et al*., 2012). A wide range of other genes influencing loss of dispersal ability have been identified in various seed and fleshy-fruited crops, and the effects and identities of many other candidate sequences are under consideration (Li & Olsen, 2016)
Fruit size	The genetic control of fruit size has been intensively researched in tomato, where the *FRUITWEIGHT2.2* (*FW2.2*) gene, which codes for a negative regulator of cell proliferation that may function as a metal cation transporter, has an important function, accounting for up to 30% of the difference in fruit weight between domesticated tomato and its wild relatives (Frary *et al*., 2000). Variation at orthologues of tomato *FW2.2* also effect fruit size in a range of other crops (Azzi *et al*., 2015). Other tomato fruit-growth-related genes have been identified, including *FW3.2* (*SlKLUH*) that encodes for a cytochrome P450 enzyme which may also play a role in regulating fruit mass in other crops (Chakrabarti *et al*., 2013)
Length of juvenile stage (unit time return to labour)	For perennial crops especially, the length of the juvenile phase of the plant is an important factor in determining labour returns. In various perennials, this has been shown to be controlled by orthologues of the arabidopsis *TERMINAL FLOWER 1* (*AfTFL1*) gene that encodes a phosphatidylethanolamine-binding protein which acts as a floral repressor (Bergonzi & Albani, 2011). Transgenic apple (*Malus domestica*) expressing *MdTFL1* antisense RNA, with reduced *MdTFL1* function, was shown to exhibit accelerated flowering (Kotoda *et al*., 2006). The use of an *Apple latent spherical virus* vector to simultaneously promote the expression of the arabidopsis *FLOWERING LOCUS T* gene and silence *MdTFL1*, through embryo inoculation immediately after germination, resulted in early flowering of the resultant apple seedlings, with the cross-pollination of these early-flowering plants producing fruits with seeds (Yamagishi *et al*., 2014).

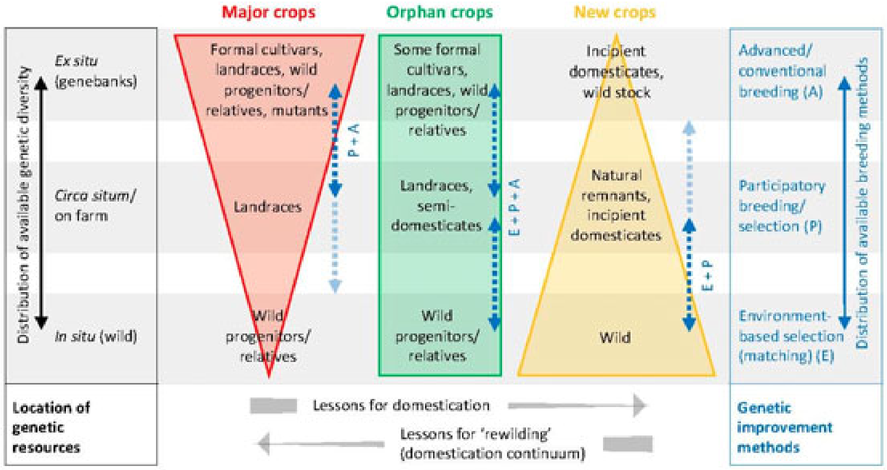

Major crops
Orphan crops
New crops
Distribution of available genetic diversity
Ex situ (genebanks)
Circa situm/ on farm
In situ (wild)
Formal cultivars, landraces, wild progenitors/ relatives, mutants
Landraces
Wild progenitors/ relatives
P + A
Some formal cultivars, landraces, wild progenitors/ relatives
Landraces, semi-domesticates
Wild progenitors/ relatives
E + P + A
Incipient domesticates, wild stock
Natural remnants, incipient domesticates
Wild
E + P
Advanced/ conventional breeding (A)
Participatory breeding/ selection (P)
Environment-based selection (matching) (E)
Distribution of available breeding methods
Location of genetic resources
Lessons for domestication
Lessons for 'rewilding' (domestication continuum)
Genetic improvement methods

Fig. 1.

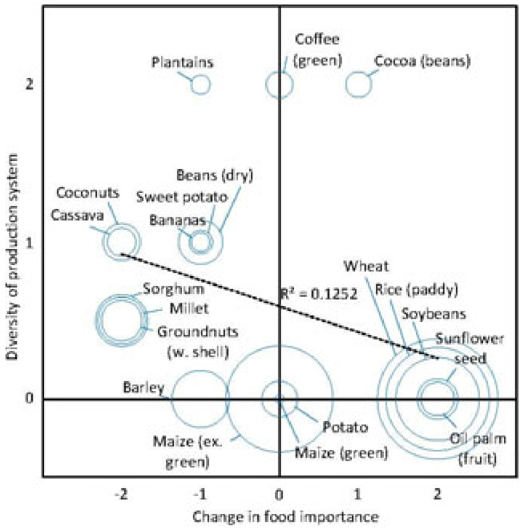

Plantains
Coffee (green)
Cocoa (beans)
Beans (dry)
Coconuts
Sweet potato
Cassava
Bananas
Wheat
R² = 0.1252
Rice (paddy)
Sorghum
Millet
Soybeans
Groundnuts (w. shell)
Sunflower seed
Barley
Potato
Maize (ex. green)
Maize (green)
Oil palm (fruit)
Diversity of production system
Change in food importance
2
1
0
-2
-1
0
1
2

Fig. 2.

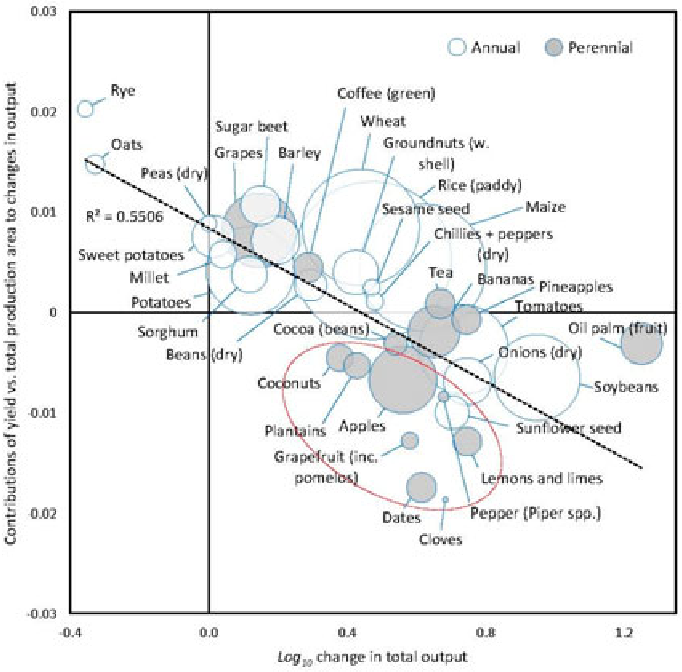

0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.4
0.0
0.4
0.8
1.2
Contributions of yield vs. total production area to changes in output
Log_{10} change in total output
○ Annual
● Perennial
R² = 0.5506
Rye
Oats
Peas (dry)
Sugar beet
Grapes
Barley
Coffee (green)
Wheat
Groundnuts (w. shell)
Rice (paddy)
Sesame seed
Maize
Chillies + peppers (dry)
Sweet potatoes
Millet
Potatoes
Sorghum
Beans (dry)
Tea
Bananas
Pineapples
Tomatoes
Cocoa (beans)
Oil palm (fruit)
Onions (dry)
Soybeans
Coconuts
Plantains
Apples
Sunflower seed
Grapefruit (inc. pomelos)
Lemons and limes
Dates
Pepper (Piper spp.)
Cloves

Fig. 3.

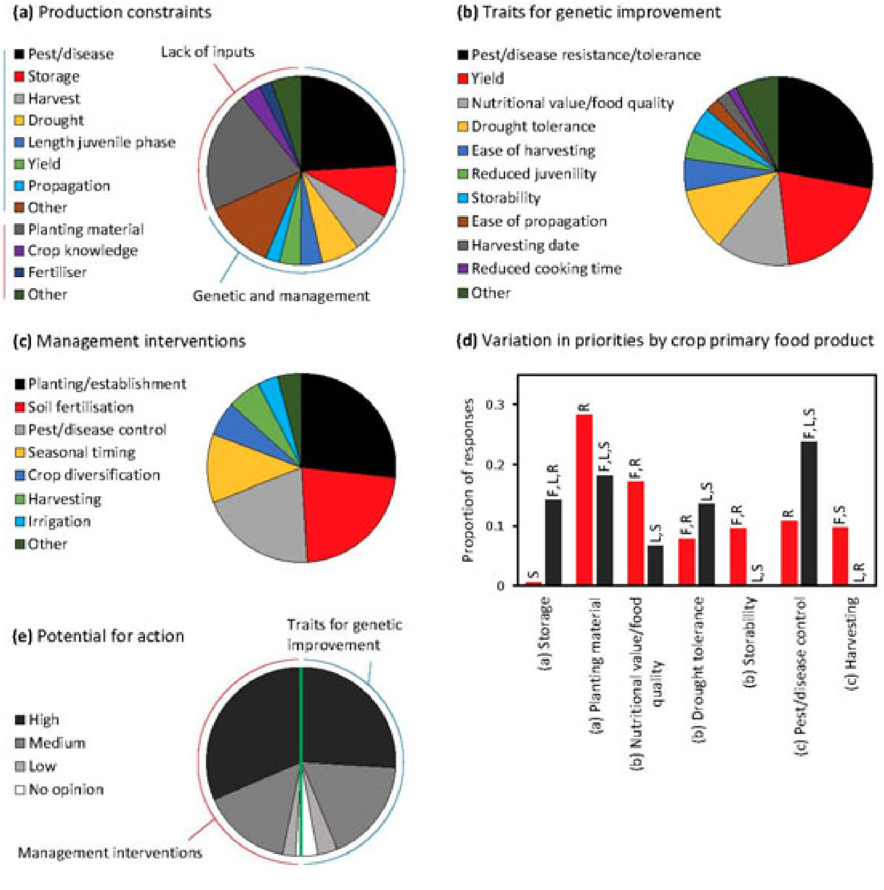
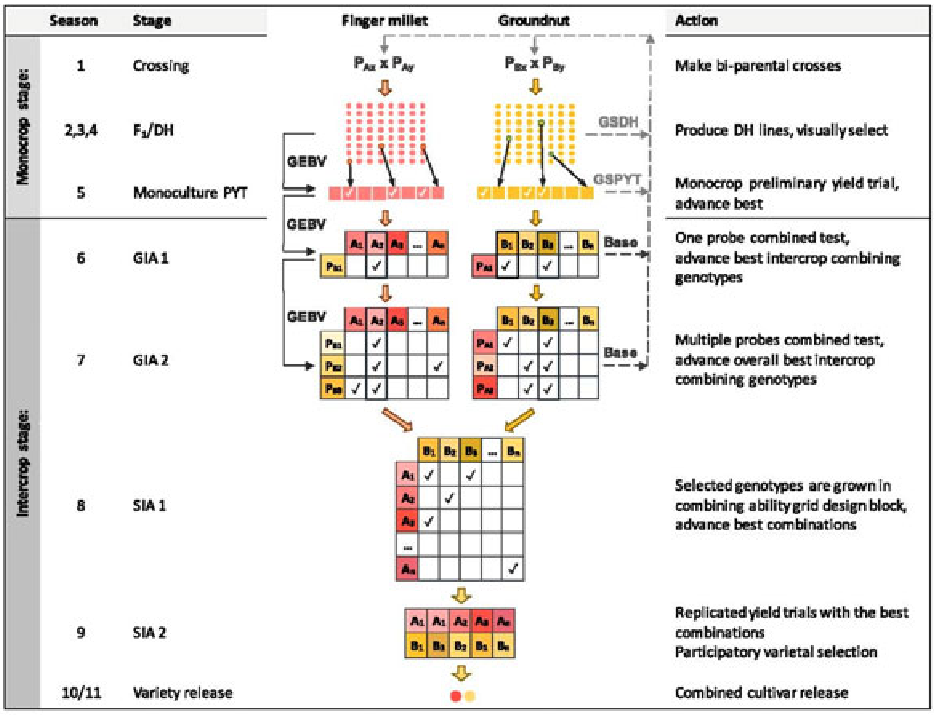

(a) Production constraints
Pest/disease
Storage
Harvest
Drought
Length juvenile phase
Yield
Propagation
Other
Planting material
Crop knowledge
Fertiliser
Other
Lack of inputs
Genetic and management
(b) Traits for genetic improvement
Pest/disease resistance/tolerance
Yield
Nutritional value/food quality
Drought tolerance
Ease of harvesting
Reduced juvenility
Storability
Ease of propagation
Harvesting date
Reduced cooking time
Other
(c) Management interventions
Planting/establishment
Soil fertilisation
Pest/disease control
Seasonal timing
Crop diversification
Harvesting
Irrigation
Other
(d) Variation in priorities by crop primary food product
Proportion of responses
0.3
0.2
0.1
0
S
F,L,R
R
F,L,S
F,R
L,S
F,R
L,S
F,R
L,S
R
F,L,S
F,S
L,R
(a) Storage
(a) Planting material
(b) Nutritional value/food quality
(b) Drought tolerance
(b) Storability
(c) Pest/disease control
(c) Harvesting
(e) Potential for action
Traits for genetic improvement
High
Medium
Low
No opinion
Management interventions

Fig. 4.



Season	Stage	Action
1	Crossing	Make bi-parental crosses
2,3,4	F_1/DH	Produce DH lines, visually select
5	Monoculture PYT	Monocrop preliminary yield trial, advance best
6	GIA 1	One probe combined test, advance best intercrop combining genotypes
7	GIA 2	Multiple probes combined test, advance overall best intercrop combining genotypes
8	SIA 1	Selected genotypes are grown in combining ability grid design block, advance best combinations
9	SIA 2	Replicated yield trials with the best combinations Participatory varietal selection
10/11	Variety release	Combined cultivar release

Fig. 5.

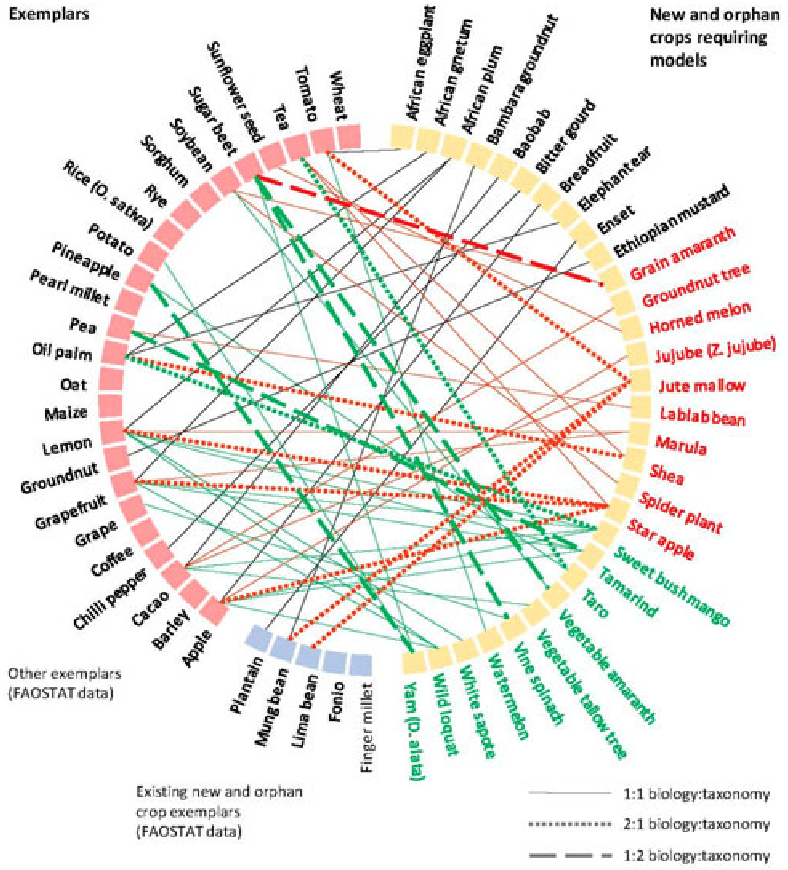

Exemplars
New and orphan crops requiring models
Wheat
Tomato
Tea
Sunflower seed
Sugar beet
Soybean
Sorghum
Rye
Rice (O. sativa)
Potato
Pineapple
Pearl millet
Pea
Oil palm
Oat
Maize
Lemon
Groundnut
Grapefruit
Grape
Coffee
Chilli pepper
Cacao
Barley
Apple
Other exemplars (FAOSTAT data)
Plantain
Mung bean
Lima bean
Fonio
Finger millet
Existing new and orphan crop exemplars (FAOSTAT data)
African eggplant
African gnetum
African plum
Bambara groundnut
Baobab
Bitter gourd
Breadfruit
Elephant ear
Enset
Ethiopian mustard
Grain amaranth
Groundnut tree
Horned melon
Jujube (Z. jujube)
Jute mallow
Lablab bean
Marula
Shea
Spider plant
Star apple
Sweet bush mango
Tamarind
Taro
Vegetable amaranth
Vegetable tallow tree
Vine spinach
Watermelon
White sapote
Wild loquat
Yam (D. alata)
1:1 biology:taxonomy
2:1 biology:taxonomy
1:2 biology:taxonomy

Fig. 6