Progress of Genomics-Driven Approaches for Sustaining Underutilized Legume Crops in the Post-Genomic Era

Uday Chand Jha1*, Harsh Nayyar2, Swarup K Parida3, Melike Bakır4, Eric J. B. von Wettberg5,6 and Kadambot H. M. Siddique7*

1ICAR-Indian Institute of Pulses Research (IIPR), Kanpur, India, 2Panjab University, Chandigarh, India, 3National Institute of Plant Genome Research (NIPGR), New Delhi, India, 4Department of Agricultural Biotechnology, Faculty of Agriculture, Erciyes University, Kayseri, Turkey, 5Plant and Soil Science and Gund Institute for the Environment, The University of Vermont, Burlington, VT, United States, 6Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia, 7The UWA Institute of Agriculture, The University of Western Australia, Perth, WA, Australia

Legume crops, belonging to the Fabaceae family, are of immense importance for sustaining global food security. Many legumes are profitable crops for smallholder farmers due to their unique ability to fix atmospheric nitrogen and their intrinsic ability to thrive on marginal land with minimum inputs and low cultivation costs. Recent progress in genomics shows promise for future genetic gains in major grain legumes. Still it remains limited in minor legumes/underutilized legumes, including adzuki bean, cluster bean, horse gram, lathyrus, red clover, urd bean, and winged bean. In the last decade, unprecedented progress in completing genome assemblies of various legume crops and resequencing efforts of large germplasm collections has helped to identify the underlying gene(s) for various traits of breeding importance for enhancing genetic gain and contributing to developing climate-resilient cultivars. This review discusses the progress of genomic resource development, including genome-wide molecular markers, key breakthroughs in genome sequencing, genetic linkage maps, and trait mapping for facilitating yield improvement in underutilized legumes. We focus on 1) the progress in genomic-assisted breeding, 2) the role of whole-genome resequencing, pangenomes for underpinning the novel genomic variants underlying trait gene(s), 3) how adaptive traits of wild underutilized legumes could be harnessed to develop climate-resilient cultivars, 4) the progress and status of functional genomics resources, deciphering the underlying trait candidate genes with putative function in underutilized legumes 5) and prospects of novel breeding technologies, such as speed breeding, genomic selection, and genome editing. We conclude the review by discussing the scope for genomic resources developed in underutilized legumes to enhance their production and play a critical role in achieving the “zero hunger” sustainable development goal by 2030 set by the United Nations.

Keywords: underutilized legumes, genomics, molecular marker, food security, transcriptomics
INTRODUCTION

Burgeoning pressure from the global human population, increasing food demands, and adverse effects of global climate change are serious concerns for global food and nutrition security (Godfray et al., 2010; Foley et al., 2011; Ebi and Loladze 2019). In addition, increasing outbreaks of plant diseases and pests, loss of arable land, and increasing environmental degradation due to excessive use of chemical fertilizers and pesticides have constrained crop yields (Godfray et al., 2010; Lesk et al., 2016). Of the various approaches for sustaining global food production without deteriorating soil and environmental health, crop diversification is needed to maintain sustainable agro-ecological systems and prevent biodiversity losses (Huñagel J. et al., 2020; Tamburini et al., 2020). Legume crops remain the third most widely grown class of crops globally (Gepts et al., 2005), providing “one third of all dietary protein nitrogen” to the human population, enriching soil fertility by fixing atmospheric nitrogen in association with symbiotically active rhizobacteria in roots (Graham and Vance, 2003), and adding rotational value to subsequent crops (Yigezu et al., 2019; Marques et al., 2020). Likewise, legume fodder and forage mitigate the rising global demand for dietary protein by livestock and provide industrial raw materials (Das and Arora 1978; Elfaki and Abdelatti 2018). Most studies have focused on major grain legumes, such as soybean, common bean, and chickpea. However, some legume crops (Supplementary Table S1) with high nutrient contents are grown in limited areas on small scales in developing countries under low input conditions and marginal land (Cullis and Kunert 2017; Kamenya et al., 2021). Despite the enormous potential of these legumes, they are neglected and known as “underutilized” legumes (Cullis and Kunert 2017; Kamenya et al., 2021). Underutilized species are rarely grown outside of a narrow geographic area, are cultivated with low chemical inputs or mechanization, are not broadly used outside of traditional cuisines, and have not been the focus of major public and private breeding companies. In the last decade, major grain legume crops have witnessed unprecedented advances in genomic resource development, including the development of reference genome sequences due to rapid advances in genome sequencing technologies, especially, next-generation sequencing (NGS). However, underutilized legume crops are lagging behind in terms of developing genomic resources. Thus, in this review we analyze the present global status of these underutilized legumes in terms of area, production, major production and nutritional quality limitation and origin (Supplementary Table S1) and discuss the available genomic resources, including their molecular marker repertoire and genome assemblies. We review the progress in genetic linkage maps and identification of trait QTLs through bi-parental mapping and genome-wide association studies of various underutilized legumes, including the downstream application of genomic assisted breeding (GAB). The discovery of various trait candidate gene(s) with putative function through transcriptome sequencing are discussed with examples. We also briefly how crop wild relatives (CWRs), whole-genome resequencing (WGRS), and pan-genome sequences could underpin novel structural variants across the whole genome in these crops. Finally, we propose the prospects and scope of novel breeding schemes—genomic selection, genome editing, and speed breeding—for enhancing genetic gain to achieve “zero hunger” in 2030.

Why Genomics and Advanced Breeding Tools for Underutilized Legumes

Underutilized legumes generally require few inputs, are rich in protein, vitamins, and minerals, and can often withstand harsh environments, including drought, extreme temperature, and waterlogging. Furthermore, these legumes replenish soil nitrogen by fixing atmospheric nitrogen through root nodules, ameliorate soil properties, and sustain agro-ecosystem services (Bhartiya et al., 2015; Ditzler et al., 2021). In addition to their role in combating nutritional and economic security, underutilized legumes play critical roles in various human diseases as they are rich in bioactive compounds and nutraceutical and medicinal properties (Prasad and Singh 2015; Bazzano et al., 2001). However, despite these benefits, there are several constraints and challenges related to the production and productivity of these legumes due to biotic and abiotic stresses (Supplementary Table S1). Furthermore, the edible seeds of some underutilized legumes contain antinutritional elements, constraining their use (Campbell et al., 1994; Tate and Ennenking 2006; Kroc et al., 2017). Combining modern genomic and traditional breeding approaches could help develop new plant types, reduce yield losses from biotic and abiotic stresses, add value for consumer preferences, and eliminate antinutritional properties.

How Minor Legumes can Catch up With Genomics

One of the aspects of the advances in DNA sequencing technology over the past two decades has been the potential to democratize research. Before the advent of next generation sequencing, performing molecular genetic research outside of a handful of species, such as fruit flies and Arabidopsis, was cost-prohibitive. Exponential declines in the cost of sequencing have made research in nearly any species not only feasible, but practical. Consequently, crops like chickpeas, pigeonpea and cowpea, once considered minor crops, now have a rich array of genome resources (e.g., Jha 2018; Varshney et al., 2021). However, there are still a range of crop species that have received less attention, due to being grown over a limited geographic extent or market demand that is mostly restricted to a small region.

For those crop species that still trail behind others for genomic resources, there is hope that lessons learned in other species can be applied to others. In legumes, where there is substantial genome synteny across the entire family [e.g., (Ren et al., 2019)], the potential for comparative genomics to speed up research in understudied species is particularly high. With improving databases for mining genomic information from more widespread cultivated legumes [e.g., (Bauchet et al., 2019; Berendzen et al., 2021)], this task has become easier than in the past.

In a range of minor legume crops, one of the foci for improvement are “domestication syndrome” traits, such as...
pod shattering, seed dormancy, seed size, and palatability. There is growing evidence that at least some of the loci controlling these traits are shared, such as for pod shattering (Ogutcen et al., 2018). With shared loci and extensive genomic synteny, either finding natural variation at these loci or using genetic modification become much easier.

**TABLE 1** Genomic resources in underutilized legumes developed during the last decade.

| Common name      | Genome size | Mapping populations | SSRs/SNPs discovered       |
|------------------|-------------|---------------------|-----------------------------|
| Adzuki bean      | 538 Mbp     | ~6                   | 7,947 EST-SSR (Chen et al., 2015a) |
| Bambara          | 550 Mbp     | ~2                   | 143,113 SSRs (Kang et al., 2015) |
| Groundnut        | —           | —                   | 1292 SSR (Chapman et al., 2015); 3,343 SNP (Uba et al., 2021) |
| Clusterbean      | 580.9 Mbp   | —                   | 5,773 (Tanwar et al., 2017); 8,687 (Patwari et al., 2017) |
| Common bean      | 1.8 Gb      | —                   | 27,066 (Ali-Quramy et al., 2019); 1,859 genomic SSRs (Tanwar et al., 2017) |
| Vetch            | —           | —                   | 5,999 SNPs and 249 InDels (Thakur and Randhawa 2018) |
| Dolichos bean    | —           | 4                   | 6,848 SSRs and 7,246 high quality SNPs (De la Rosa et al., 2020) |
| or hyacinth bean | —           | —                   | 9,320 DA-T seq based SNPs and 15,719 SilicoDart markers |
| Grasspea bean    | 8.2 Gb      | ~2                   | 651,827 SSRs and 288 SSRs (Yang et al., 2014); 3,204 EST-SSR (Hao et al., 2017); 146,406 SNPs (Hao et al., 2017) |
| Horse gram       | 400 Mbp     | —                   | 6,915 SSRs (Bhardwaj et al., 2013) |
| Lima bean        | 622 Mbp/1 C | —                   | 3,942 SNPs (Mahesh et al., 2021) |
| Mothbean         | —           | 1                   | 10,497 SNPs (Garcia et al., 2021) |
| Mungbean         | 494–555 Mb  | ~19                  | 13,134 EST-SSRs (Chen et al., 2015b); and 200,808 SSRs in mungbean (Kang et al., 2014); 775,831 high-confidence SNPs (Kang et al., 2014) |
| Narrow-leafed lupin | 924 Mbp     | ~9                   | 8,966 SNPs (Hao et al., 2021); 233,799 SNPs (Bangar et al., 2021) |
| Red clover       | 420 Mbp     | ~3                   | 469,907 SNPs (Kosztrewicz et al., 2017) |
| Ricebean         | 414 Mbp     | ~2                   | 3,942,975 SNP (Li et al., 2019) |
| White lupin      | 451 Mbp     | 1                   | 1,621 genic-SSR and 1844 SNPs (Raizada et al., 2019) |
| Yellow lupin     | —           | 1                   | 5,765 SNPs (Somta et al., 2019) |
| Urd bean         | 574 Mbp     | 3                   | 1853 SSRs (Chapman et al., 2015); 12,956 SSRs (Vatanparast et al., 2016) |
| Winged bean      | 1.22 Gbp/C  | —                   | 5,190 SNPs (Vatanparast et al., 2016) |
| Zombi pea        | —           | 2                   | 4,044,822 SNPs in zombiopsis (Amkul et al., 2019) |

Advances in Genomic Resource Development in Underutilized Legumes

In the last decade, rapid advances in genome sequencing technologies have enriched the genomic resources, including genome-wide distributed high-throughput molecular markers especially, simple sequence repeats (SSRs) and single
nucleotide polymorphisms (SNPs), transcriptomes, and whole-genome assemblies, of various underutilized legumes.

**Molecular Marker Resources**

Hybridization-based molecular markers, such as restriction fragment length polymorphisms (RFLP), and PCR-based molecular markers, such as RAPD, SSR markers, have been used to analyze, tag, and map trait gene(s) in various underutilized legumes (Bohra et al., 2014). However, the arrival of next-generation sequencing technology (NGS) based high-throughput (HTP) markers, especially SNPs, has replaced traditional PCR-based molecular markers for genotyping.

| Crop name | Genotype | Pubmed ID | Chromosome no. | Size of genome | No. of protein coding genes | Genome coverage | Sequencing platform used | References |
|-----------|----------|-----------|----------------|----------------|-----------------------------|----------------|--------------------------|------------|
| Cyamopsis tetragonoloba (L.) Taub. | Vaviloskij 130 | — | — | 1.2 Gb | — | 5x | (Illumina and Oxford Nanopore) | Grigoreva et al. (2019) |
| Cyamopsis tetragonoloba (L.) Taub. | RGC-936 | — | — | 550.31 Mbp | 34880 | 366.73x | (Illumina, 10x Chromium and Oxford Nanopore) | Gaikwad et al. (2020) |
| L. purpureus | — | 30535374 | 2n = 22 | 395.47 Mb | 20,946 | — | HiSeq 2000 platform | Chang et al. (2019) |
| Lupinus albus | AMiGA | 31980615 | 2n = 50 | 451 Mb | 38258 | 164x | PacBio Sequel platform | Bárbara Hufnagel et al. (2020) |
| Lupinus angustifolius | Tanjil | 27557478 | 2n = 40 | 609 Mb | 33,076 | 162.8x | Illumina | Wang P. et al. (2021) |
| Lupinus angustifolius | Tanjil | 33249667 | 2n = 40 | 615.8 Mb | 33907 | 156x | PacBio Sequel II platform | Sangasawa et al. (2021a) |
| Macrotyloma uniflorum | HPK-4 | — | — | 259.2 Mb | — | — | Illumina HiSeq 2000 | Mahesh et al. (2021) |
| Medicago polymorpha | Huaiyang Jinhuaiae | 33642569 | — | 1.2 Gb | — | — | Illumina, PacBio and Hi-C technologies | Cui et al. (2021) |
| M. ruthenica | — | — | — | 904.13 Mb | 50,162 | — | PacBio, Illumina, 10xGenomics, and Hi-C | Mou Yin et al. (2021) |
| Narrow leafed lupin | Tanjil | 23734219 | — | 279.1 Mb | 24,521 | — | Illumina HiSeq platforms | Jenniches et al. (2021) |
| Narrow leafed lupin | Tanjil | — | — | 538 | 57,807 | 27x | Illumina HiSeq platforms | Yang et al. (2013a) |
| Phaseolus lunatus L. | G27455 | 33514713 | — | 512 Mbp | 28,326 | 10x | Illumina HiSeq | Garcia et al. (2021) |
| Phaseolus acutifolius A. Gray | Frijol Bayo | — | 2n = 22 | 684 Mb | 27,538 | 101.28x | Illumina HiSeq platforms | Moghadam et al. (2021) |
| Phaseolus acutifolius A. Gray | wild accession | — | 2n = 22 | 676 Mb | 27,096 | — | Illumina HiSeq platforms | Yang et al. (2013a) |
| Red clover | Tatra | 24500806 | — | 314.6 | 47,398 | 50x | Illumina HiSeq 2000 | Isolánek et al. (2014) |
| Red clover | MUVAS B | 26617401 | — | 309 Mb | 40,868 | 30x | Illumina HiSeq 2000 | De Vega et al. (2015) |
| T. subterraneum L. | Daliak | 27545089 | — | 471.8 Mb | 42,706 | — | Illumina HiSeq and HiSeq 2000 | Hirakawa et al. (2016) |
| T. subterraneum L. | TSUd_r1.1 | 28111887 | — | 512 Mb | 31,272 | 341x | Illumina HiSeq 2000, GS FLX + | Kaur et al. (2017) |
| Vigna radiata | VC 1973A | 25384727 | 2n = 22 | 579 Mb | 22,472 | — | Illumina HiSeq 2000, GS FLX + | Yang et al. (2014) |
| Vigna angularis | Jingnong 6 | 34275211 | 2n = 22 | 475 Mb | 30,958 | — | PacBio RS II platform | Ha et al. (2021) |
| Vigna angularis var. angularis | IT213134 | 25626881 | 2n = 22 | 612 Mb | 26,857 | — | Illumina HiSeq 2000 | Kang et al. (2015) |
| Vicia sativa | KSR5 | 26617401 | — | 1.5 Gb | 31,146 | 146x | HiSeq2000 | Shirsasawa et al. (2021) |
| Vigna mungo | Pant U-31 | — | 2n = 22 | 475 Mb | 18655 | — | Illumina and Nanopore sequencing | Jegadeesan et al. (2021) |
| Vigna mungo | Chai Nat 80 | — | 2n = 22 | 499 Mb | 29,411 | 21.72x | Illumina HiSeq x Ten | Pootakham et al. (2021) |
| V. subterranea | — | 30535374 | — | 535.05 Mb | 31,707 | — | HiSeq 2000 platform | Chang et al. (2019) |

**TABLE 2** | List of genome sequence assembly of underutilized legume crops.
Second- and third-generation sequencing technologies have enabled the mining of massive numbers of SSRs and SNPs through whole-genome sequencing, WGRS, and transcriptome sequencing efforts in various crops, including underutilized legumes (Edwards and Batley 2010).

Likewise, the advent of NGS-based HTP genotyping platforms, such as Illumina’s GoldenGate assay, Illumina’s HiSeq 4000 platform, and Illumina’s Infinitium SNP array, enabled the discovery of copious SNPs across multiple genomes that facilitate a range of investigations, including the diversity of genebank collections (Sokolkova et al., 2020). Aiming at comprehensive mining of SSR markers for Vigna species including cowpea, mungbean and adzuki bean, microsatellite database VigSatDB has been developed (Jasrotia et al., 2019). A comprehensive list of molecular markers, mapping populations available in various underutilized legumes are in Table 1. Thus, these molecular markers will provide the foundation for implementing genomic assisted breeding for improving genetic gain in underutilized grain legumes.

**De Novo Genome Sequencing of Underutilized Legumes**

Adzuki bean (Vigna angularis var. angularis) (2n = 2x = 22) is an important grain legume of Asiatic origin (Kang et al., 2015). The draft genome sequence of adzuki bean was assembled on 11 pseudo-chromosomes, estimating 612 Mb or 75% of the estimated genome and high-confidence 26,857 protein-coding genes (Kang et al., 2015) (Table 2). Yang K. et al. (2015) assembled a draft genome assembly of “Jingnong 6” cultivar covering 450 Mb of the total genome.

Bambara groundnut (Vigna subterranea) (2n = 2x = 22) is an important legume crop, rich in protein (18–26%), carbohydrate (63%), and fat (6.5%) and having inherent drought tolerance capacity (Shegro et al., 2013). It originated from West Africa and is mainly grown in sub-Saharan areas, especially Nigeria (Olkologu et al., 2012). Chang et al. (2019) assembled the genome sequence of bambara groundnut, with a genome size of ~535.05 Mb with 31,707 protein-coding genes.

Mungbean (Vigna radiata, 2n = 2X = 22) is a warm-season legume crop, originated from India and mostly grown in South and Southeast Asian countries. Kang et al. (2014) first assembled the mungbean genome sequence, estimating 421 Mb or 80% of the total genome size and 22,427 protein-coding genes, with scaffold length 431 Mb and N50 length 35.4 Mb covering 314 Mb. Recently, a mungbean genome sequence was assembled with a scaffold length 431 Mb and N50 length 35.4 Mb covering 314 Mb. Recently, a mungbean genome sequence was assembled with a scaffold length 431 Mb and N50 length 35.4 Mb covering 314 Mb.

Urdbean (Vigna mungo, 2n = 2x = 22), native to Indian subcontinent, mostly grown in South and Southeast Asian countries (Kaewwongwal et al., 2015), is a rich source of dietary protein, vitamins, folate, and iron (Kakati et al., 2010). The genome assembly of Chai Nat 80 cultivar measured 499 Mb with an N50 length of 5.2 Mb (Pootakham et al., 2021). Subsequently, Jegadeesan et al. (2021) assembled a genome assembly of urdbean, measuring 475 Mb or 82% of the genome with scaffold N50 of 1.42 Mb and 42,115 genes with coding sequence.

Cluster bean (Cyamopsis tetragonoloba, 2n = 2x = 14), native to west Africa and India, an important commercial legume crop widely grown in India and parts of Africa, contains hetero-polysaccharide called guar gum or galactomannan used extensively in the cosmetic and pharmaceutical industries (Gillett 1958). Gaikwad et al. (2020) assembled the first genome sequence of RGC-936 cultivar, measuring 550.31 Mb with N50 length of 78.27 Mbps and 34,680 protein-coding genes.

Dolichos bean (Lablab purpureus) (2n = 2x = 22) is a versatile legume crop of African origin, rich in seed protein and highly tolerant to various abiotic stresses (Maass et al., 2010). It is mostly cultivated in tropical and sub-tropical regions globally (Maass et al., 2010). The genome assembly of Lablab purpureus was constructed recently, with an estimated 395.47 Mb genome size and 20,946 protein-coding genes (Chang et al., 2019).

Grass pea (Lathyrus sativus) is a climate-resilient legume of Central Asia and Abyssinia origin, diploid (2n = 2x = 14), cool-season legume species (Kamphuis et al., 2015; Emmrich et al., 2020) primarily grown on the Indian subcontinent and in northern and eastern Africa, including Ethiopia (Kumar et al., 2011). The assembled genome size of Elv1 was measured at 8.12 Gbp with scaffold N50 value of 59.7 kbp and 33,819 high-confidence genes (Kamphuis et al., 2015; Emmrich et al., 2020).

Horsegram [Macrotyloma uniflorum (Lam.) Verdc.], native to tropical southern Asia, is a diploid legume (2n = 20, 22) grown in India, Africa, and Australia (Arora and Chandel 1972). The genome sequence of the HPK-4 genotype was assembled on ten pseudomolecules measuring 259.2 Mb or 89% of the total length of the assembled sequence (Shirasawa et al., 2021a). Another genome assembly of accession PHG-9, measuring 279.1 Mb with 24,521 annotated genes has recently been constructed (Mahesh et al., 2021).

Red clover (Trifolium pratense L.; Fabaceae, 2n = 2x = 14) is an important forage legume of European origin, with a genome size of 418 Mbp. Isótovánek et al. (2014) completed a de novo assembly of the red clover genome, comprising ~314.6 Mbp.

Likewise, subterranean clover presumed to be originated from Southern Australia, belonging to Trifolium genus, is an annual diploid (2n = 2x = 16) pasture legume with 540 Mbps genome size (Kaur et al., 2017). Hirakawa et al. (2016) assembled the genome sequence of T. subterraneum L., measuring 471.8 Mb or 85.4% of the whole genome and containing 42,706 protein-coding genes. Subsequently, Kaur et al. (2017) assembled an advanced genome assembly of T. subterraneum L., estimating 512 Mb with 31,272 protein-coding genes.

Tepary bean (Phaseolus acutifolius A. Gray), native to the Sonoran Desert and a sister species of common bean, is gaining attention due to its inherent capacity for biotic and abiotic stress tolerance (Moghaddam et al., 2021) and important source traits for improving biotic and abiotic stress tolerance in common bean (Moghaddam et al., 2021). A reference genome assembly of cultivated landrace Frijol Bayo, possessing inherent heat tolerance, was constructed using Illumina X10 and HiSeq platforms and PACBIO with 101.28x sequence coverage, and...
measured 512,626,114 bp with 27,538 high-confidence genes (Mohgaddam et al., 2021).

White lupin (Lupinus albus L. 2n = 50) originated from Mediterranean region, contains high protein content (30–40% whole seed) (Bähr et al., 2014) and can use higher soil phosphorus than other legume crops due to its special “cluster root” structure (Lambers et al., 2013). However, improving yield stability and minimizing anti-nutritional alkaloids in white lupin seed through conventional breeding remains challenging. Hence, to elucidate the function of various trait gene(s) related to quality and quantitative importance, Bárbara Hufnagel et al. (2020) assembled a high-quality genome sequence of white lupin, scaling 451 Mb and 38,258 annotated protein-coding genes.

Likewise, narrow-leafed lupin (Lupinus angustifolius) is an important grain legume of Mediterranean origin, enriched with dietary protein (40–45%) and fiber (25–30%) (Lee et al., 2006). Hane et al. (2017) assembled the draft genome sequence of Tanjil cultivar, estimating 609 Mb and 33,076 protein-coding genes. Subsequently, Wang et al. (2021a) constructed an improved cultivar, estimating 609 Mb and 33,076 protein-coding genes. Hane et al. (2017) assembled the draft genome sequence of Tanjil, measuring 615.8 Mb with contig N50 = 5.65 Mb, using a long-read whole-genome sequencing approach.

Common vetch (Vicia sativa, 2n = 14) originated from Near Eastern centre of diversity, is a wild and partially domesticated legume crop with a genome size of 1.8 Gb (Shirasawa et al., 2021b). It is used as silage and hay for livestock feeding. The reference genome assembly has been assembled, spanning 1.5 Gb and 31,146 genes (Shirasawa et al., 2021b).

### Quantitative Trait Mapping Through Bi-parental and Multi-Parental Schemes

As most of the traits with agricultural importance including biotic, abiotic stress tolerance and quality traits are governed by multiple gene(s)/quantitative trait loci (QTL). In order to map these traits various molecular breeding approaches are available to breeders, including family based bi-parental mapping approach, marker-assisted backcrossing. Subsequently, the availability of high-throughput molecular markers has accelerated the precise mapping of various trait QTLs through employing novel molecular breeding schemes including MutMap, multi-parental cross (MAGIC), genome-wide association mapping, genomic selection and QTL seq approach (Meuwissen et al., 2001; Cavanagh et al., 2008; Takagi et al., 2013; Takagi et al., 2015). In underutilized legumes several bi-parental mapping populations based on interspecific and intraspecific crosses have been developed aiming at constructing genetic linkage map and mapping/tagging targeted trait QTLs of agronomic importance (for details Table 3). However, mapping resolution of detected QTLs through bi-parental mapping approach remains low.

| Crop | Mapping population | Type of population | Size of LG map | Number of marker/loci assigned | Marker density | References |
|------|--------------------|--------------------|----------------|-------------------------------|---------------|------------|
| Vigna radiata | Vigna radiata × V. umbellata | RIL | 1,291.7 cM | 538 SNPs | 2.40 cM | Mathivathana et al. (2019) |
| Lupinus angustifolius L | 83A-476 × P27255 | RIL | 2,599 cM | 34,574 markers/3,508 loci | 0.67 cM | Liu et al. (2016a) |
| Yellow lupin | Wodjil cultivar × P28213 | RIL | 2,261.3 cM | 2,458 loci | 2.29 cM | Iqbal et al. (2019) |
| Lupinus angustifolius L | Emir × LAE-1 | RIL | 3,042 cM | 4602 markers | 2.18 cM | Garcia et al. (2021) |
| Lupinus angustifolius L | 83A-476 × P27255 | RIL | 2,500.8 cM | 9,792 loci | 0.85 cM | Hane et al. (2017) |
| Phaseolus lunatus L | UC 92 × UC Haskell | RIL | 1,064 cM | 522 loci | 2.18 cM | Garcia et al. (2021) |
| Horsegram | HPK-4 × HPKM-193 | F2 | 980 cM | 1,263 SNPs | — | Taylor et al. (2021) |
| Vigna vexillata | TVNu 240 × TVNu 1,623 | F2 | 1,740.9 cM | 6,529 loci | 0.27 cM | Amkul et al. (2019) |
| Vigna vexillata | Dahuaye × Jiyu 9-1 | RIL | 1,060.2 cM | 1,946 bin markers | 0.54 | Wang et al. (2020) |
| Vigna angustifolius | Vigna angustifolius × V. angustifolius var. nipponensis | F2 | 1,365.0 cM | 2,904 loci | 0.47 cM | Wang et al. (2021b) |
| Vigna aconitifolia | TN67 × ICPMO056 | F2 | 1,016.8 cM | 172 loci | 7.34 cM | Yundeng et al. (2019) |
| Horsegram | HPK4 × HPKM249 | RIL | 1,423.4 cM | 211 loci | 9.6 cM | Chahota et al. (2020) |
| Vigna radiata | UC 1973A × V2984 | RIL | 1,365.0 cM | 2,904 loci | 0.47 cM | Wang et al. (2021b) |
| Lathyrus | BEGE00277 × BEGE023542 | F2 | 724.2 cM | 307 loci | 2.4 cM | Santos et al. (2018) |
| Vigna vexillata | V. vexillata (JP235863) × | F2 | 704.8 cM | 262 loci | 2.6 cM | Dachapak et al. (2018) |
| — | wild V. vexillata (AusTRCF6614) | — | — | — | — | — |
| Bambara groundnut | ITA688 × Ankpa4 | F2 | 1,395.2 cM | 223 markers | — | — |
| — | Tiga Nicuru × DipC | F2 | 1,376.7 cM | 293 markers | — | — |

**References**
- Isobe et al. (2009)
- Kozak et al. (2019)
- Hane et al. (2017)
- Taylor et al. (2021)
- Garcia et al. (2021)
- Shirasawa et al. (2021a)
- Wang et al. (2020)
- Wang et al. (2021b)
- Yundeng et al. (2019)
- Chahota et al. (2020)
- Kang et al. (2014)
- Santos et al. (2018)
- Dachapak et al. (2018)
- Ho et al. (2017)
| Crop | Trait | Mapping population | QTL | LG group | Type of marker | PV% | References |
|------|-------|--------------------|-----|----------|----------------|-----|------------|
| Bambara groundnut | Internode length | IITA686 x Anikpa4, F₂ 263 | One major QTL | LG2 | DAR/Tseq markers | 33.4 | Ho et al. (2017) |
| | | Tiga Nicuru x DipC, F₃ 71 | | | | | |
| Lupinus angustifolius | Gray leaf spot | 83A:476 x P2755, F₈ RIL | One major QTL, LOC109334326 | LG19 | Microsatellite fragment | 98 | Zhou et al. (2021) |
| | | | LOC109334327 | | | | |
| Lupinus angustifolius | Vernalisation | Chittick × Geenburg, F₂ and RIL | elf, Trimethylguanosine | LG14 | SNP | 81.95% | Taylor et al. (2021) |
| | | | | | | | |
| Lupinus albus | Anthracnose | KievP27174 F₈, RIL | antr04_1, antr05_1, antr04_2, antr05_2 | ALB02, ALB04 | SNP | 14.6-25 | Rychel-Bielska et al. (2020) |
| | | | | | | | |
| Lupinus luteus | Domestication related traits | Wodjil×P28213, RIL(156) | Vernalisation responsiveness locus, Alkaloid content, flower and seed | YL-21, YL-06 | SNP, presence of Absence variation Marker | 83% | Iqbal et al. (2020) |
| | | | | | | | |
| Lupinus luteus | Anthracnose resistance | AluProt-CGNA × Pt385149 | Anthracnose resistance QTL | 4, 10, 11, 13, 23 | SNP | 75-83% | Lichtin et al. (2020) |
| | | | | | | | |
| Lupinus luteus | and early flowering | F₂ (188) | Days to flowering QTL | qDFW01, qDFW02, qDTM01 | LG1,4,6 | SSR, RAPD, COS | 7.3-55.3% | Chahota et al. (2020) |
| Macrotyloma | Drought and yield | HPK4 × HPKM249 (RIL,190) | | | | | |
| | | | | | | | |
| Vigna radiata | Drought | VC 1973A × V2984, RIL, 187 | Height4-1, Height5-1 | LG4, 5 | SNP | 6.2-30 | Ha et al. (2021) |
| | | | | | | | |
| Vigna radiata | Flower initiation | — | F₁₁₁₁, F₁₀₁₁ | LG4,9 | SNP | 6.4-24 | Ha et al. (2021) |
| | | | | | | | |
| Vigna radiata | No. of branches | — | Branch3-1 | LG3 | SNP | 6.4 | Ha et al. (2021) |
| | | | | | | | |
| Vigna radiata | No. of nodes | — | Node4-1, Node11-1 | LG4, 11 | SNP | 6.3-20 | Ha et al. (2021) |
| | | | | | | | |
| Vigna radiata | Synchronous maturity | — | SPMA-1, SPMB-1 | LG4, 7 | SNP | 6.8-10.3 | Ha et al. (2021) |
| Vigna aconitifolia | C. chinensis | TN67× IPCMO066, F₂(188) | qVacBr2.1 and qVacBr5.1 | LG2 and 5 | SSR | — | Somta et al. (2018) |
| Vigna mungo | C. maculatusresistance. | BC48 × TC2210, RIL(150) | qCm_PDS2.1, qCm_AUDPS6.1, qCm_AUDPS7.1, qVmunBr6.1 and qVmunBr6.2 | LG2, 6 and 7 | SNP | 7.28-30% | Somta et al. (2019) |
| | | | | | | | |
| Vigna radiata | Indented Leaflet | Dahuahe × Jilyu 9-1 | Indented Leaflet QTL | LG3 and LG10 | SNP | 39.70% | Wang et al. (2020) |
| | | | | | | | |
| Vigna radiata | C. chinensis | — | VΠGP1 and VΠGP2 | LG5 | — | and 45.4% | Zhang et al. (2021) |
| Vigna angularis | Flowering time | Vigna nipponensis: Yesheng10 × JinHong9218 | Fld2.2 and Fld3.2, Fld5.1 vs. Fld5., and Fld5.2 vs. Fld5.5 | LG03, LG05 | SLAF | 66-71% | Mao-Sen Liu et al. (2016) |
| Vigna angularis | Seed size | Vigna angularis × V. angularis var. nipponensis | 12 seed size related QTLs | LG2, 4, 5,6 and 9 | Indels | 3-22% | Wang et al. (2021b) |

(Continued on following page)
Therefore, to increase the resolution of trait QTLs novel breeding scheme viz., genome-wide association study (GWAS), nested association mapping and MAGIC has been developed. We believe these approaches could be implemented in underutilized legumes to increase the resolution of trait QTLs.

**Progress in High-Density Genetic Map Development for Trait Quantitative Trait Loci Discovery and Mapping**

Initially, morphological-based markers, isozymes, RFLP, amplified fragment length polymorphisms (AFLP), randomly amplified polymorphic DNAs (RAPD), and SSR markers were used to construct preliminary genetic linkage maps in various underutilized legumes [for details, (Bohra et al., 2014)]. However, the increasing ease of developing high-throughput SNP markers derived by GBS, restriction site-associated DNA sequencing (RAD-seq), and whole genome resequencing has facilitated developing highly dense/saturated consensus linkage maps in various underutilized legumes.

Several genetic maps of mungbean based on SSR markers have been developed (Bohra et al., 2014). Later, a genetic map measuring 1,060.2 cM was developed from an intraspecific mapping population (Wang et al., 2020) and a denser genetic map with 1,291.7 cM and harboring 538 SNPs was developed from an interspecific mapping population derived from Vigna radiata × V. umbellate cross (Mathivathana et al., 2019) (Table 4).

A comprehensive genetic map of red bean (V. mungo) covering 1,588.7 cM with 3,675 SNPs was developed (Somta et al., 2019). Based on a F₂ population, Kai Yang et al. (2015) developed an initial genetic map in adzuki bean measuring 1,031.17 cM. Wang et al. (2021b) presented a denser genetic map measuring 1,365.0 cM in adzuki bean (V. angularis). In zombi pea (V. vexillata), a high-density linkage map spanning 1740.9 cM harboring 6,529 SNPs with an average distance of 0.27 cM between markers was developed from an F₂ mapping population of TVNu 240 × “TVNu 1,623” (Amkul et al., 2019).

Hane et al. (2017) presented a high-density linkage map of narrow-leafed lupin measuring 2,500.8 cM with 9,972 loci and Iqbal et al. (2019) developed a high-density linkage map of yellow lupin measuring 2,261.3 cM. Santos et al. (2018) developed a genetic map of lathyrus covering 724.2 cM with 307 loci. Chahota et al. (2020) presented a genetic map for horse gram measuring 1,423.4 cM with 211 loci (Table 4).

The above linkage maps can be used to identify various traits of biotic, abiotic stress tolerance, agronomic, and culinary importance in numerous underutilized legumes. The selected major trait QTLs identified in the last decade based on bi-parental mapping populations are listed in (Table 4). Biotic stress remains the most significant yield stress in underutilized grain legumes globally. The increased availability of genomic resources, especially molecular markers, has identified/tagged various disease-resistant QTLs/gene(s); for example, one major QTL qCm_PD6.1 against Callosobruchus chinensis (bean weevil) and another QTL qCm_PD6.1 against Callosobruchus maculatus (cowpea weevil) have been identified (Amkul et al., 2019). Likewise, four major QTLs (antr04_1, antr05_1, antr04_2 and antr05_2) controlling anthracnose resistance explaining 14–25% (Rychel-Bielska et al., 2020) of the phenotypic variation in white lupin. Restriction site-associated DNA sequencing derived SNP markers were used as candidate markers for the R gene of phomopsis stem blight disease resistance in narrow-leafed lupin (Yang et al., 2013b). Recently, one major QTL with LOC109334325, LOC109334327 underlying candidate genes was deciphered for gray leaf spot disease in narrow-leafed lupin (Zhou et al., 2021).

Like biotic stresses, abiotic stresses, particularly drought, causes significant yield losses in underutilized legumes (Liu et al., 2017; Chahota et al., 2020). Several QTLs contributing to drought tolerance have been discovered in mungbean (Liu et al., 2017), and horse gram (Chahota et al., 2020).

**TABLE 4 | (Continued) List of selected QTLs identified in various underutilised legume crops.**

| Crop                | Trait                                | Mapping population | QTL                | LG group  | Type of marker | PV%  | References                |
|---------------------|--------------------------------------|--------------------|-------------------|-----------|----------------|------|--------------------------|
| *Vigna aconitifolia*| Domestication related traits          | TN87 × ICPMO056 F₂(188) | Domestication related trait | —         | SSR, RAD-seq | 5.9–52% | Yundeng et al. (2019)    |
| *Vigna vexillata*   | Domestication related traits          | JP235863 x AusTRCF66514 F₂(139) | Domestication related trait | 37 QTLs related to | LGs 5, 6, 7, 8, 10 and 11 | SSR, RAD-seq | 52% | Dachapak et al. (2018)  |
| *Vigna vexillata*   | 22 domestication-related traits      | V. vexillata (JP235863) x wild V. vexillata (AusTRCF66514) | Domestication related trait | 37 QTLs | LG1, 2, 3, 4, 5, 6, 7, 8, 9 | SSR, RAD-seq | — | Dachapak et al. (2018)  |
| *Vigna vexillata*   | C. chinensis resistance              | TVNu 240 × TVNu 1,623 F₂(139) | One major and three minor QTLs | —         | SNP            | — | Amkul et al. (2019)    |
| *Vigna vexillata*   | C. maculatus resistance              | —                   | One major and one minor QTLs for C. maculatus | —         | —              | — | —                       |

Frontiers in Genetics | www.frontiersin.org 8 April 2022 | Volume 13 | Article 831656
Low seed-alkaloid content (<0.02%) is a prime objective of quality improvement in lupin. In lupin the *iucundus* allele is a major gene regulating seed alkaloid content. Several mapping populations have been developed for identifying low alkaloid controlling QTLs and gene(s). Li et al. (2011) identified a microsatellite-anchored fragment length polymorphism-derived PCR marker (*LucLi*) linked to the low-alkaloid locus *iucundus* (0.9 cM). Likewise, Lin et al. (2009) developed a sequence-specific PCR marker (PauperM1) closely linked (1.4 cM) to the low-alkaloid locus *pauper* in white lupin (*Lupinus albus* L.). Moreover, of five SNP markers co-segregating the *pauper* locus in a set of 140 lupin accessions, the *LAGI01_35805_F1_R1* marker was highly linked with this gene and could be used in low seed alkaloid lupin breeding programs (Rychel and Książkiewicz, 2019). Subsequently, Kroc et al. (2019) developed a co-dominant derived cleaved amplified polymorphic sequence (dCAPS) marker (*iuc_RAP2-7*) from the RAP2-7 candidate gene of alkaloid locus *iucundus* responsible for seed alkaloid content in narrow-leaved lupin, which could be used in marker-assisted breeding for low alkaloid content in lupin. Furthermore, fine mapping of this seed alkaloid controlling genomic region unveiled four candidate gene(s)—LOC109339893, LOC109339862, LOC109339875 and LOC109339876—on LG7 in the interval of 20.70–20.89 Mb (Wang et al., 2021a).

**Genome-Wide Association Study Approach for Trait Quantitative Trait Loci Identification With Increased Resolution**

GWAS is gaining popularity for uncovering genotype–phenotype associations in various plant species, including underutilized legumes (Huang and Han 2014; Liu and Yan 2019), by establishing the genetic basis of the genotype–phenotype association for the trait of interest in a large panel of diverse accessions based on multiple crossing-over events over the recent demographic history of a taxa (Huang and Han 2014). Due to the unprecedented advances in NGS technology, an increasing repertoire of HTP markers in several underutilized legumes have helped to identify loci associated with aspects of complex trait architecture. GWAS has been assisted by the subsequent availability of genome-wide SNP markers for various traits, including phenological traits, quality/nutritional traits, biotic and abiotic stresses, and yield and yield-related traits, in many underutilized legumes (Plewniński et al., 2020). In narrow-leaved lupin, a GWAS incorporating massive analysis of cDNA ends (MACE) markers in 126 genotypes uncovered significant MTAs related to flower initiation, maturity, plant height, and yield traits (Plewniński et al., 2020). The underlying candidate genes were *Lup019134, Lup015264, Lup021911*, and *Lup021909* for flower initiation, *Lup015264* and *Lup004734* for maturity, *Medtr1g030750* for plant height, and *Lup021835* and *Lup022535* for yield traits (Plewniński et al., 2020).

GWAS has been used increasingly for dissecting complex QTLs controlling various abiotic stresses in crop plants, including underutilized legumes. To elucidate the underlying genomic regions attributing macro- and micro-nutrients in mungbean seeds, Wu et al. (2020) identified 43 MTAs related to calcium, iron, manganese, phosphorus, sulfur, and zinc using inductively coupled plasma (ICP) spectroscopy and GBS-derived SNPs in a set of 95 global mungbean accessions. The explained phenotypic variation ranged from 1 to 38%. Further, Reddy et al. (2021) used a GBS-based GWAS study to dissect the molecular basis of phosphorus uptake efficiency and phosphorus utilization efficiency in 120 mungbean genotypes. The authors uncovered 116 SNPs in 61 protein-coding genes related to phosphorus uptake efficiency and phosphorus utilization efficiency traits. The significantly associated SNPs explained phenotypic variation ranging from 17 to 20% for total phosphorus utilization (under low phosphorus) and it ranged from 15 to 21% for phosphorus utilization efficiency. Six candidate genes—VRADI01G04370, VRADI05G20860, VRADI06G12490, VRADI08G20910, VRADI08G00070 and VRADI09G09030—regulating phosphorus uptake efficiency and phosphorus utilization efficiency were deciphered (Reddy et al., 2021).

Recently, recruiting 5,041 SNPs in a minicore collection of 293 mungbean accessions identified four significant MTAs for maturation and hypocotyl color within the *Vradi02g04380* gene on chromosome 2 encoding zinc finger A20 and AN1 domain stress-associated protein (Sokolkova et al., 2020). Despite the popularity of GWAS for elucidating marker-trait associations, it has some drawbacks regarding population structure and low-frequency causal alleles causing false negative results (Korte and Farlow 2013). To minimize and overcome the population structure related problems, artiﬁcially designed populations such as MAGIC and nested association mapping, could be used [for details (Alseck et al., 2021)].

**Crop Wild Relatives and Their Genome Assembly for Exploring Novel Trait Genes in Underutilized Legumes**

CWRs, including those of underutilized legumes, are a hidden reservoir of novel trait gene(s), offering scope for broadening genetic diversity in crop breeding programs (Warschefsky et al., 2014; Zhang and Batley, 2020). In the past, during domestication process, several genes associated with adaptive traits conferring abiotic stress tolerance were lost rendering modern cultivated crop plants to adapt poorly under stress condition (Warschefsky et al., 2014; Zhang and Batley, 2020). However, CWRs serve as reservoir of these biotic and abiotic stress adaptive genes. Thus, recapturing these genes from CWRs through introgression and novel breeding tools could facilitate in increasing the fitness of genepool (Burgarella et al., 2019). Several CWRs of underutilized legumes. e.g., *V. nakashimae*, are potential sources of bruchid resistance (Somta et al., 2006) and salinity tolerance (Yoshida et al., 2016) in adzuki bean. Likewise, harnessing bruchid resistance genes/genomic regions from *Vigna radiata* var. *sublobata* can improve bruchid resistance in mungbean (Schafleitner et al., 2016) ([Table 5](#table5)). In urd bean, *V. mungo* var. *silvestris* could be promising for transferring bruchid and mungbean yellow mosaic India virus resistance genes into high-yielding urd bean breeding lines (Sourframani et al., 2006; Souframanien et al., 2010). Further, the
genomic sequences of wild underutilized legumes have been assembled to gain insight into the novel trait genes of CWRs. Whole-genome sequencing of *M. ruthenica* offered novel insights into many genes, including the FY3/FAR1 gene family conferring higher drought tolerance in cultivated *M. sativa* (Wang et al., 2021c). Mou Yin et al. (2021) advocated evidence for multiple family genes and TF family genes, viz., C2H2, CAMTA and NAC attributing various abiotic stress tolerances through chromosome-scale genome sequencing of *M. ruthenica*. Novel SNP and InDel markers were recovered from genome sequencing of *V. radiata* var. *sublobata*; the wild relative accession TC1966 of mungbean could be useful for exploring biotic and abiotic stress tolerant genomic regions through comparative mapping of cultivated mung bean (Kang et al., 2014). Thus, these CWR genomic resources could be used to develop climate-resilient grain legume cultivars.

**Implications of Genomic-Assisted Breeding in Underutilized Legumes**

Current advances in genomic resource development in underutilized legumes have enabled breeders to develop improved cultivars. For example, tagging various traits in narrow-leaved lupin, such as LanFTc1 PCR-based INDEL markers for vernalization responsiveness locus *KuJulius* (Nelson et al., 2017; Plewiński et al., 2019; Taylor et al., 2019), InDel2, InDel10, and PhtjM7 for *PhtjR* (Yang et al., 2013b; Yang H. et al., 2015), Anseq3 and Anseq4 for *Lamr1* (Yang et al., 2012), and TP222136 and TP47110 markers for antr04_1/antr05_1 and TP338761 for antr04_2/antr05_2 (anthracnose resistance) (Rychel-Bielska et al., 2020), the iucLi co-dominant marker (Li et al., 2011) and RAP2-7 PCR-based dCAPS marker for major alkaloid content locus *iucundus* (Kroc et al., 2019) are available. Likewise, a diagnostic marker LAGI01_35805_F1_R1 linked to *pauper* locus controlling low alkaloid content in white lupin could be used for practicing MAS of white lupin lines with low-alkaloid content (Rychel and Księżek 2019). Moreover, co-dominant markers linked to the *tardus* (Li et al., 2010) and *lentus* (Li et al., 2012) genes, attributed to low pod shattering, could be of interest for developing zero shattering narrow-leaved lupin using marker-assisted breeding.

Similarly, CEDG261 and DMB-SSR160 markers linked to bruchid resistance could be used in GAB in moth bean breeding programs (Somta et al., 2018). Downstream application of GAB in concert with other novel breeding approaches for enhancing genetic gain in various underutilized legumes is depicted in Figure 1.

**Transcriptomics Resources as a Component of Functional Genomics for Gene Discovery With Function in Underutilized Legumes**

The advent of NGS-based RNA-seq technology assessing global gene expression has offered a platform for the discovery of functional markers, including EST-SSRs and SNPs, capturing gene space and shedding light on a myriad of trait candidate genes and their plausible functions (O’Rourke et al., 2013; Yang et al., 2017; Glazińska et al., 2019). Previously, EST markers, microarrays, and cDNA libraries were the major functional genomic resources for investigating the function of various trait genes. For example, cDNA library sequencing identified 125,821 unique sequences (O’Rourke et al., 2013) in white lupin.

Subsequently, advances in transcriptome sequencing facilitated the discovery of many unigenes and differentially expressed genes for various traits of importance for details (see Table 6). Transcriptome studies have also shed light on the functional role of various underlying candidate gene(s) controlling seed biology, plant phenology, biotic and abiotic stress tolerance, yield traits, and nutritional quality traits, including alkaloid regulation in narrow-leaved lupin, β-N-oxalyl-L-α, β-diaminopropionic acid (β-ODAP) in grass pea and condensed tannin in winged bean (Kroc et al., 2019; Yang et al., 2017; Xu et al., 2018).

---

**Table 5** List of CWRs source of novel trait gene in various underutilized legumes.

| Crop       | Wild species                                      | Importance                                    | References                  |
|------------|--------------------------------------------------|-----------------------------------------------|-----------------------------|
| Adzuki bean| *V. nakashimae*                                   | Bruchid resistance                            | Somta et al. (2006)         |
| Adzuki bean| *V. angulans* var. *nipponensis*                  | Domestication- and fitness-related traits     | Kaga et al. (2005)          |
| Adzuki bean| [JP205833 of *V.nakashimae*](#)                   | Salinity tolerance                            | Yoshida et al. (2016)       |
| Grasspea   | *L. articulatus* L. ([IG64792 and IG65197](#))    | *Orobanche crenata*                           | —                           |
|            | *L. aphaca* L. and *L. ochrus*                    | *O. foetida* Poir                             | —                           |
| Mungbean   | JP 2118749                                       | Bruchid resistance and domestication related traits | —                            |
| Mungbean   | *Vigna radiata* var. *sublobata*                  | Bruchid resistance                            | Isemura et al. (2012)       |
| Mungbean   | *Vigna radiata* var. *sublobata* TC1966           | Bruchid resistance                            | Kaewwongwai et al. (2015)   |
| Mungbean   | *Vigna umbellata*                                 | *Mungbean yellow mosaic virus*                 | Schafleitner et al. (2016)  |
| Wild vigne | *V. rukkuensis*, *V. trilobata*, *V. vexillata*   | Salinity tolerance                            | Sudha et al. (2015)         |
| Urn bean   | *V. luteola*, *V. manihia*                        | Bruchid resistance                            | Yoshida et al. (2020)       |
| —          | *V. mungo* var. *silvestris*                      | *Mungbean yellow mosaic virus*                 | —                           |
| —          | —                                                 | *India virus (MVMV)*                          | Souframanien and Gopalakrishna (2006) |
| —          | —                                                 | —                                             | Souframanien et al. (2010)  |

---

Frontiers in Genetics | www.frontiersin.org 10 April 2022 | Volume 13 | Article 831656
In association with small RNA sequencing, degradome sequencing and transcriptome sequencing helped unravel key molecular players, including various phytohormones and metabolic pathways involved in floral development and organ abscission of *L. luteus* (Glazinska et al., 2017; Glazińska et al., 2019). Moreover, participation of small RNA related to seed biology and the conglutin gene encoding seed storage protein was demonstrated in a transcriptome study in narrow-leafed lupin (DeBoer et al., 2019).

Transcriptome studies could improve our understanding of the regulatory mechanisms of the complex network of gene(s), pathogenesis-related genes, phytohormone signaling response, and non-coding RNAs mediating plant immune responses to attacking pathogens (Almeida et al., 2014; Dasgupta et al., 2021). To gain insight into the molecular mechanisms involved in conferring rust resistance in grasspea, an RNA-seq study in rust-responsive grasspea (resistant vs. susceptible) revealed the upregulation of salicylic acid and abscisic acid in the rust-resistant genotype and downregulation of jasmonate and ethylene pathways in the susceptible genotype (Almeida et al., 2014) (Table 6). Additionally, several pathogenesis-related genes and the mildew resistance locus O (MLO)-like resistance gene were discovered in this study.

An RNA-seq study offered insight into the participatory role of WRKY, NAC and MYB transcription factors, phytoene synthase, cytochrome P450, and JAZ and LOX genes attributing to mungbean yellow mosaic virus (MYMV) resistance (Dasgupta et al., 2021).

Likewise, transcriptome studies can decipher the complex molecular mechanisms and underlying possible candidate gene(s) networks during perceiving abiotic stress signaling and mediate various abiotic stress tolerances by activating antioxidant mechanisms and other cellular protective mechanisms, enabling plants to acclimate to abiotic stress (Bhardwaj et al., 2013; Butsayawarapat et al., 2019; De la Rosa et al., 2020).

A de novo transcriptome analysis of two contrasting horse gram genotypes for drought tolerance revealed the involvement of various TFs (NAC, MYB, and WRKY families) in conferring drought stress tolerance (Bhardwaj et al., 2013). De novo transcriptome sequencing of contrasting drought tolerant and sensitive genotypes of common vetch revealed a plethora of differentially expressed genes under water stress (De la Rosa et al., 2020). Most of the genes mediating drought tolerance are associated with cell wall modification, oxidative stress response and ABA response (De la Rosa et al., 2020). In zombi pea, a comparative transcriptome analysis revealed up-regulatory activity of glycolysis and fermentative genes in the waterlogging-sensitive genotype; in contrast, the waterlogging-tolerant genotype had enhanced activity of auxin-regulated lateral root initiation, aquaporin, and peroxidase genes (Butsayawarapat et al., 2019) (Table 6).

Deciphering the underlying genes and molecular function of quality parameters, including nutritional and industrially important parameters, using transcriptomic studies could improve these traits (Yang et al., 2017, Xu et al., 2018; Tyagi et al., 2018). Small RNA sequencing indicated the involvement of several miRNAs and their target genes coding for carbohydrate metabolism, kinase, and enzymes for regulating galactomannan biosynthesis in cluster bean (Tyagi et al., 2018) (Table 6). The authors also discovered two novel unigenes, mannosyltransferase/mannan synthase (ManS) and UDP-
| Crop            | Trait                  | Candidate genes/Unigenes/DEG | Function                                      | Platform used         | References                  |
|-----------------|------------------------|------------------------------|-----------------------------------------------|-----------------------|-----------------------------|
| Common vetch    | Drought stress         | 2,646 transcripts are DEG    | Redox homeostasis, cell wall modifications    | Illumina HiSeq 2,500  | De la Rosa et al. (2020)    |
|                 | tolerance              |                              |                                               |                       |                             |
| Common vetch    | Pod shattering         | 1,285 DEGs and 575 upregulated unigenes | Hydrolyase activity                          | HiSeq 2000            | Dong et al. (2017)          |
|                 |                        | 710 downregulated unigenes   |                                               |                       |                             |
| Guar            | Root development       | 102,479 unigenes             | Root development                              | Illumina HiSeq 2,500  | Thakur and Randhawa (2016)  |
|                 |                        |                              |                                               |                       |                             |
|                 | stress tolerance       |                              | Stress tolerance                              |                       | Tanwar et al. (2017)       |
| Guar            |                        | 11,308                       | Carbohydrate, protein, lipid, energy          | Illumina HiSeq        |                             |
|                 |                        |                              |                                               |                       |                             |
| Guar            | Galactomannan          | 187 known and 171 novel miRNAs differentially expressed | Regulating galactomannan pathway | Illumina NextSeq 500  | Tyagi et al. (2018)         |
|                 |                        |                              |                                               |                       |                             |
| Guar            | Biosynthesis           |                              |                                               |                       |                             |
| Guar            |                        | 38423 DEGs                   | Metabolic process, cellular process           | Illumina              | Rawal et al. (2017)         |
| Guar            |                        | 61,508 putative genes        | Biological process, cellular component and molecular function | Illumina HiSeq 2,500  | Al-Qurainy et al. (2019)    |
| Guar            | Galactomannan          |                              | Galactomannan biosynthesis pathway            | Illumina Hiseq. 4000  | Chaudhry et al. (2019)      |
|                 | Biosynthesis           |                              |                                               |                       |                             |
| Horsegram       | Drought                | 21,887 unigenes              | Calmodulin binding factor, heat shock protein | Illumina GAIIx        | Bhardwaj et al. (2013)     |
| Lathyrus        | Rust tolerance         | 134,914 contigs              | Regulating phytohormone signalling            | Illumina Hiseq2000    | Almeida et al. (2014)       |
| Lathyrus        | Acochlya isthryi       | 738 units                    | Cell wall metabolism                          | DeepSuperSAGE         | Almeida et al. (2015)       |
| Lathyrus        | Rust                   | 27,431 unigenes              |                                               | Illumina NextSeqTM 500| Hao et al. (2017)           |
| Lathyrus        | β-ODAP                 | 4620 and 3,498 contigs down regulated | Hormone metabolism, cell wall degradation     | Illumina Hiseq2000    | Santos et al. (2018)       |
| Lathyrus        |                        | 213,258 unigenes             | Carbohydrate and                              | Illlumina-HiSeq 3,000 | Xu et al. (2018)            |
|                 |                        |                              | sulfur assimilation/metabolism                |                       |                             |
| L. angustifolius| quinozidine alkaloids  | 12 candidate genes, RAP2-7, AP2/ERF TF | Quinozidine synthesis                        | Illumina HiSeq 1,500  | Kroc et al. (2019)          |
| L. angustifolius| quinozidine alkaloids  | 33 genes related to lupin alkaloid biosynthesis | Copper amine oxidase                         | Illumina HiSeq 2,500  | Yang et al. (2017)          |
| Lupinus albica  | Phosphorus             | 2,128 sequences differentially expressed in response to Pi deficiency | Cluster root development                      | Illumina GA-IIx       | O’Rourke et al. (2015)     |
| Lupinus         |                        | 10,240 transcripts           | Peroxidase and anthocyanin biosynthesis       | Illumina HiSeq 2000   | Kamphuis et al. (2015)     |
| angustifolius   |                        |                              |                                               |                       |                             |
| Medicago        | Drought tolerance      | 3,905 genes and 50 mRNAs     | gma-miR171-5p and mtr-miR396a-5p down regulated | Illumina Hiseq4000    | Shi et al. (2021)           |
| Mungbean        | MyMV                   | 1881, 1,449, 1,583 and 1,140 genes as up-regulated | Defence related activity                     | Illumina HiSeq 2,500  | Dasgupta et al. (2021)     |
|                 |                        | 1,423, 1,154, 1,396 and 1,152 genes as down-regulated |                                               |                       |                             |
| Mungbean        | Osmotic response       | 13 OSCA genes                |                                               |                       | Mou Yin et al. (2021)      |
|                 |                        | 1,245                        |                                               |                       | Changyu Liu et al. (2016)  |

(Continued on following page)
| Crop               | Trait                   | Candidate genes/Unigenes/DEG | Function                                           | Platform used        | References       |
|--------------------|-------------------------|-----------------------------|---------------------------------------------------|----------------------|------------------|
| Psophocarpus       | —                       | 5,053 transcript have predicted functions | biological process, cellular component | Illumina platform    | Vatanparast et al. (2016) |
| Trifolium tetragonolobus | —             | —                           | betaine aldehyde dehydrogenase                   | —                    | —                |
| Trifolium ambiguum | Rhizome development     | 276 DEGs involved in hormone signaling and transduction | rhizome growth and development | PacBio sequencing    | Yin et al. (2020) |
| Trifolium pratense | Drought                 | 45181 contigs               | Role of proline, malate and pinitol              | Illumina sequencing  | —                |
| Trifolium pratense | Seed setting            | 1,196 DEGs                 | Contributing to drought tolerance                | Illumina sequencing  | —                |
| Trifolium pratense | Regrowth                | —                           | These gene(s) involved in seed setting           | Illumina sequencing  | —                |
| Trifolium pratense | Iso-flavonoid           | 143 iso-flavonoid synthesis genes | Role various genes and long non coding RNAs contributing to iso-flavonoid synthesis | Illumina HiSeq X Ten platform | Shi et al. (2021) |
| Trifolium pratense | Flower pigmentation     | 6,282 DEGs, CHS, F3′H, F3′5′H, UFGT, FLS, LAR, ANS, and DFR | Anthocyanin flavonoid biosynthetic pathway and flavonoid biosynthetic pathway | Illumina Hiseq X10 | Heshan Zhang et al. (2018) |
| Urd bean           | —                       | 2,306 DEGs                 | Cytochrome c-type biogenesis protein             | Illumina MiSeq       | Raizada and Souframanien (2019) |
| Urd bean           | —                       | —                           | DnaJ protein homolog 1                           | —                    | —                |
| Urdbean            | —                       | 29564 transcript contigs    | Uncharac- terized protein LOC108329961            | Illumina             | Souframanien and Reddy (2015) |
| Vigna angularis    | —                       | 65,950 unigenes            | RING-H2 finger protein                           | —                    | —                |
| Vigna angularis    | —                       | —                           | A serine/threonine protein kinase                | —                    | —                |
| Vigna angularis    | —                       | —                           | Lipase ROG1-like protein                         | —                    | —                |
| Vigna angularis    | Drought                 | 324,219 and 280,056 transcripts | Hormone signal transduction                      | Illumina HiSeq 2000 | Jo et al. (2016) |
| Vigna angularis    | —                       | 5,337 DEGs                 | Transcript or translation processes              | Illumina HiSeqX      | Zhu et al. (2020) |
| Vigna angularis    | —                       | —                           | Ubiquitin proteasome system                      | —                    | —                |
| Vigna angularis    | Heat stress             | Uregulation of PIP1-1 and PIP2-7 in leaves and the TIP2-1 | Induction of aquaporin genes                     | qRT-PCR              | Qi et al. (2021) |
| Trifolium repens   | Water logging           | 982 and 1,133 DEGs         | Induction of Cell wall modification               | Illumina HiSeq 4000  | Butsayawarapat et al. (2019) |
| —                  | —                       | —                           | Acquaporin, and peroxidase genes                 | —                    | —                |
| —                  | —                       | —                           | Auxin Metabolism                                 | —                    | —                |
D-glucose 4-epimerase (UGE), targeted by Ct-miR3130, Ct-miR3135, and Ct-miR3157 miRNAs. Likewise, an RNA-seq study revealed preferential expression of 2,535 and 2,724 genes in endosperm and 3,720 and 2,530 genes in the embryo involved in guar gum biosynthesis (Hu et al., 2019).

Transcriptome assembly through RNA-seq identified several candidate genes regulating quinolizidine alkaloids (QAs) biosynthesis, an anti-nutritional factor in narrow-leaved lupin (Kamphuis et al., 2015; Yang et al., 2017; Kroc et al., 2019). Short-read sequencing using Illumina HiSeq2500 in association with long-read sequencing using PacBio technology of high QA-containing genotypes identified 33 candidate genes associated with QA biosynthesis in narrow-leaved lupin (Yang et al., 2017). Furthermore, transcriptome profiling offered insight into the genes involved in the accumulation and degradation of β-N-oxalyl-L-α, β-diaminopropionic acid (β-ODAP), a neurotoxin found in grasspea (Xu et al., 2018). Similarly, RNA-seq analysis of high- and low-tannin-containing lines of winged bean, using Illumina Nextseq 500, revealed 1,235 differentially expressed contigs in these two lines. Several genes related to condensed tannin were elucidated, including anthocyanidin 3-O-glucosyltransferase (A-3GOT), anthocyanidin synthase (ANS), chalcone synthase (CHS) phenylalanine ammonia-lyase (PAL) (Singh et al., 2017).

Scope of Genomic Selection/Genomic Prediction for Increasing Genetic Gain in Underutilized Legumes

The decoding of various underutilized legume genome sequences and resequencing efforts have made SNP markers accessible, providing great opportunities to perform genomic selection (GS). This approach has been used for estimating the genomic breeding value of tested individuals without any prior phenotypic information by measuring the genome-wide marker effect based on various prediction models (Meuwissen et al., 2001). Thus, the benefits of GS could be harnessed for the selection of progenies with known genotypic scores with high “genetic merit” for improving genetic gain.

Assessing anthracnose resistance in white lupin using GS based on GBS-derived SNPs in the ridge regression BLUP model, Rychel-Bielska et al. (2020) reported a moderately high predictive ability (0.56). Application of GS is very limited in minor legumes; however, increasing repertoire of genome wide SNP markers will greatly assist in implementing GS for improving future genetic gain in these legumes.

Scope of Speed Breeding, an Innovative Approach for Enhancing Breeding Efficiency in Underutilized Legumes

Speed breeding could be used to increase breeding efficiency by shortening the breeding cycle and reducing plant space, cost, and labor resources, thereby increasing genetic gain (Watson et al., 2018; Hickey et al., 2019). Speed breeding protocols have been established by optimizing photoperiod, daylength, and temperature in various legume crops, including soybean (Fang et al., 2021), chickpea (Samineni et al., 2019), pigeonpea (Saxena et al., 2019), and pea (Mobini and Warkentin, 2016). However, this approach has not been implemented in any underutilized legumes. Thus, the establishment of a speed breeding protocol could open up new avenues for improving genetic gain in various underutilized legumes more quickly than traditional breeding methods.

Resequencing and Pangene Assembly for Capturing Novel Structural Variations Across the Whole Genome

With the declining costs of genome assembly construction, whole genome resequencing is gaining popularity for uncovering genomic regions controlling traits of agronomic importance in a large set of global crop germplasm (Hufnagel B. et al., 2020). The WGRS approach can elucidate the causal candidate gene(s)/genomic regions associated with traits of interest. Like other major grain legumes, WGRS has been used in underutilized legume crops (Yang H. et al., 2015; Hufnagel B. et al., 2020). The resequencing of nine lupin cultivars discovered 180,596–795,735 SNP markers and 243 candidate diagnostic markers linked to the PhlR (phomopsis stem blight disease) gene (Yang H. et al., 2015). Of these candidate diagnostic markers, nine were validated in commercial cultivars, offering an opportunity to practice marker-assisted breeding for phomopsis stem blight disease resistance in narrow-leaved lupin.

Resequencing 11 modern cultivars, two landraces, and one wild relative of white lupin and comparing them with the reference genome sequence revealed the recent breeding history of white lupin (Hufnagel B. et al., 2020). Similarly, 38 narrow-leaved lupin accessions, including 19 wild and 19 cultivated types, with 19x coverage of the genome were resequenced to reveal the genomic signal for domestication and genes associated with the domestication process (Wang et al., 2021a). A selective sweep analysis in the same study identified 303 genomic regions under strong selection, with 8.2% of the genome under selection associated with domestication. Further, these selective sweeps harbored nine key domestication-related traits, including early flowering, reduced pod shattering, white flower, and low alkaloid (Wang et al., 2021a). WGRS efforts of three mungbean accessions using the Ion Torrent Personal Genome Machine™ (PGM™) platform identified 233,799 SNPs and 9,544 insertions and deletions in coding and non-coding regions, revealing great opportunity for future mung bean improvement using genomic-assisted breeding (Bangar et al., 2021).

Previously, molecular biologists and geneticists have relied mainly on the “single reference genome sequence” of a species for genomic analyses within and across species (Sherman and Salzberg, 2020; Della Coletta et al., 2021). However, the single reference genome sequence does not explain all of the genomic variation/structural variants available within and across species; “pangenomics” studies can capture all of the genomic information in a species. The pangenome refers to the entire non-redundant DNA sequences existing in a species, constituting the “core” genome common to all individuals in a species, with
“dispensable” genome the variable fraction or “accessory” genome (Tettelin et al., 2005; Sherman and Salzberg, 2020; Della Coletta et al., 2021; Lei et al., 2021). In the context, Hufnagel et al. (2021) constructed the pangenome of white lupin using a “map to pan” approach (Hu et al., 2017) by sequencing 39 accessions, which identified 32,068 core genes and 14,822 dispensable genes. They also identified 333 selection sweeps related to low alkaloid content and candidate genes (LaDHDPs, LaHLT, and LaAT) controlling alkaloid content. Pangenome analyses of other underutilized legumes could provide novel insights into genomic variation for future trait discovery.

Several legume genera have multiple domesticated species. For example, Vigna has 10 domesticated taxa, Phaseolus seven, and Lupinus four. Super pan-genomes across these genera might have immense power to provide insight into similarities in domestication syndromes, the genetic basis of traits governing geographic distribution, and disease and pest resistance.

Hope and Progress of Genome Editing in Underutilized Grain Legumes

Despite the success of transferring gene(s) of interest into high-yielding cultivars, environmental biosafety and regulatory governing bodies have not allowed the widespread adoption of transgenic technology (Zhang Y. et al., 2018).

Genome editing tools, especially the CRISPR/Cas9 based technique, has revolutionized functional genomics and plant breeding, creating novel genetic variation in plants by editing targeted genes of interest with precision and efficiency (Chen and Gao, 2014). Examples of genome editing in various crops are increasing (Chen and Gao 2014; Zhang Y. et al., 2018); however, there has been limited success in legume species. Notable instances of CRISPR/Cas9 mediated genome editing have been reported in soybean (Cai et al., 2015; Sun et al., 2015; Han et al., 2019), cowpea (Ji et al., 2019) and Medicago truncatula (Michno et al., 2015). In case of cowpea, Ji et al. (2019) employed CRISPR/Cas9 based genome editing tool in the symbiosis receptor -like kinase target gene VnSYM RK that controls nodule symbiosis in cowpea. The edited plant exhibited complete inhibition in nodule formation and consequently, the mutant plants were unable to synthesise nodules in association with Sinor hizobium sp. strain NGR234. Furthermore, complete male and female sterile plants were generated by editing SPOI1-1 gene through CRISPR/Cas9 technology in cowpea (Juranič et al., 2020). In the context of underutilized legume, the CRISPR/Cas9 genome engineering technique was used to edit the isoﬂavone synthase gene contributing to rhizobial defense signaling in red clover (Dinkins et al., 2021). Furthermore, gene-editing technology in association with base editors and prime-editing could be harnessed for de novo domestication of CWRs of underutilized legumes and “reengineering of metabolism” to increase resilience and enhance nutritious value (Gasparini et al., 2021; Nasti and Voytas 2021).

Scope of de Novo Domestication of Underutilized Legumes

Crop wild relatives are the richest reservoir of genetic diversity for improving various biotic and abiotic stress resistance in crop plants and could therefore be used as new crops through “de novo domestication” or “redomestication” process (Fernie and Yan 2019; Von Wettberg et al., 2021). Domestication of new legume underutilized crops from their wild relatives could strengthen crop diversity, and thus be vital for sustainable agriculture (Zhang et al., 2018). Among the various underutilized grain legume species, Vigna stipulacea could be targeted for de novo domestication due to its inherent capacity for drought and salinity stress tolerance and reduced pod shattering (Takahashi et al., 2019). Likewise, being an incompletely domesticated species and having inherent stress tolerance ability against biotic and abiotic stress, hairy vetch (Vicia villosa) is an ideal legume crop for de novo domestication (Renzi et al., 2020).

Of the various approaches, mutagenesis and forward screening and CRISPR/Cas9 based gene editing are important techniques for introducing domestication-related traits in wild relatives for de novo domestication (Shapter et al., 2013; Li et al., 2018). Ethyl methanesulfonate mutagenesis and forward screening enabled the domestication of Vigna stipulacea Kunzke by selecting mutants with reduced pod shattering and reduced seed dormancy (Takahashi et al., 2019). Likewise, CRISPR/Cas9 genome editing technology could be used to eliminate g-glutamyl-b-cyano-alanine (GBCA) toxin from seeds of common vetch (Vicia sativa), providing a zero-toxin vetch variety for combating the rising global protein demand (Nguyen et al., 2020).

CONCLUSION AND FUTURE PERSPECTIVES

Given the rising demand for food, feed, and forage, there is an urgent need to develop sustainable food resources. Underutilized legumes are versatile crops with great potential for mitigating global food security challenges, but they are lagging behind major legumes in terms of genomic resource development. More genomic sequencing of CWRs, landraces, and improved breeding lines will provide novel insights into genomic variations for investigating evolution, domestication events, and the diversification of underutilized legumes. Increasing genomic resources will allow increased genome-assisted breeding of these legumes. Likewise, WGRS in association with GWAS and pangenome integration with GWAS could underpin the causal genes/haplotypes of complex traits of interest. Emerging genome editing techniques could play a critical role in minimizing toxins or negative parameters associated with various nutritional quality traits, such as editing GBCA encoding gene(s) in common vetch, BOAA encoding gene(s) in grasspea, and genes involved in producing QAs in white lupin. These technologies also have great potential...
for introducing de novo domestication in CWRs by removing phenotypically undesired traits in various CWRs of underutilized legumes. Moreover, genomic selection and speed breeding approaches could enhance genetic gain in underutilized legumes. The rich diversity in these underutilized legumes needs proper collection, conservation, and characterization (Kamenya et al., 2021). Furthermore, the establishment of sound varietal releases and seed distribution systems could play a central role in popularizing these climate-smart underutilized legumes among farmers (Bohra et al., 2020). Disseminating knowledge on the global demand and profitability of these legumes needs strengthening via extension services, especially in developing countries (Kamenya et al., 2021). Hence, collective genomics, novel breeding knowledge, and sound seed system approaches could improve underutilized legume productivity for securing global food security.

AUTHOR CONTRIBUTIONS

UJ and HN developed the conceptual structure. UJ and KS prepared the original draft. EvW, SP, and MB contributed specific sections. KS edited manuscript. All authors read and approved the final manuscript.

FUNDING

EvW is partially supported in work on underutilized legumes by USDA Hatch funding and by the Ministry of Science and Higher Education of the Russian Federation as part of World-class Research Center program: Advanced Digital Technologies (Contract No. 075-15-2020-934 dated 17 November 2020).

ACKNOWLEDGMENTS

UJ acknowledges support from Indian Council of Agricultural Research (ICAR), New Delhi.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.831656/full#supplementary-material

REFERENCES

Abdallah, F., Kumar, S., Amri, A., Mentag, R., Kehel, Z., Mejri, R. K., et al. (2021). Wild Lathyrus Species as a Great Source of Resistance for Introgenesis into Cultivated Grass Pea (Lathyrus sativus L.) against Broomrape Weeds (Orobanche crenata Forsk. and Orobanche foetida Poir.). Crop Sci. 61, 263–276. doi:10.1002/csc2.20399

Adegbeyea, T. T., Abberton, M. T., AbdelGadir, A. H., Dianda, M., Maziya-Dixon, B., Oyotami, O. A., et al. (2019). Nutrient and Antinutrient Composition of Winged Bean (Psophocarpus tetragonolobus (L.) DC.) Seeds and Tubers. J. Food Qual. 2019, 1–8. doi:10.1155/2019/3075208

Al-Qurainy, F., Alshameri, A., Gaafar, A. R., Khan, S., Nadeem, M., Alameri, A. A., et al. (2019). Comprehensive Stress-Based De Novo Transcriptome Assembly and Annotation of Guar (Cyanopsis tetragonoloba (L.) Taub): An Important Industrial and Forage Crop. Int. J. Genomics 2019, 7295859. doi:10.1155/2019/7295859

Almeida, N. F., Krezdorn, N., Rotter, B. r., Winter, P., Rubiales, D., and Vaz Patto, M. C. (2015). Lathyrus Sativus Transcriptome Resistance Response to Ascophyta lathyrI Investigated by deepSuperSAGE Analysis. Front. Plant Sci. 6, 178. doi:10.3389/fpls.2015.00178

Almeida, N. F., Letão, S. T., Krezdorn, N., Rotter, B., Winter, P., Rubiales, D., et al. (2014). Allelic Diversity in the Transcriptomes of Contrasting Rust-Infected Genotypes of Lathyrus sativus, a Lasting Resource for Smart Breeding. BMC Plant Biol. 14 (1), 1–15. doi:10.1186/s12870-014-0376-2

Alseekh, S., Kostova, D., Bulut, M., and Fernie, A. R. (2020). Genome-wide Association Studies: Assessing Trait Characteristics in Model and Crop Plants. Cell. Mol. Life Sci. 78, 5743–5754. doi:10.1007/s00018-021-03868-w

Amarowicz, R., Estrella, I., Hernández, T., and Troszyńska, A. (2008). Antioxidant Activity of Extract of Azuki Bean and its Fractions. J. Food Lipids 15 (1), 119–136. doi:10.1080/09714755.2007.100116.x

Amkul, K., Wang, L., Somta, P., Wang, S., and Cheng, X. (2019). Construction of a High Density Linkage Map and Genome Dissection of Bruchid Resistance in Zongbi Pea (Vigna vexillata (L.) A. Rich.). Sci. Rep. 9 (1), 1–10. doi:10.1038/s41598-019-48239-5

Arora, R. K., and Chandel, K. P. S. (1972). Botanical Source Area of Wild Herbage Legumes in India. Trop. Grassland 6, 213–221.
Huang, X., and Han, B. (2014). Natural Variations and Genome-wide Association Studies in Crop Plants. Annu. Rev. Plant Biol. 65, 531–551. doi:10.1146/annurev-plant-052013-035715

Hufnagel, B., Marques, A., Soriano, A., Marquès, L., Divol, F., Doumas, P., et al. (2020). High-quality Genome sequence of white Lupin Provides insight into Soil Exploration and Seed Quality. Nat. Commun. 11, 1–2. doi:10.1038/s41467-019-14197-9

Hufnagel, B., Soriano, A., Taylor, J., Divol, F., Kroc, M., Sanders, H., et al. (2021). Pengenome of white Lupin Provides Insights into the Diversity of the Species. bioRxiv 2021 Jan 1. doi:10.1101/2021.01.16.243373

Hufnagel, J., Reckling, M., and Ewert, F. (2020). Diverse Approaches to Crop Diversification in Agricultural Research. A Review. Agron. Sustain Dev. 40, 1–17. doi:10.1007/s11593-020-00617-4

Hupeghe, C. (1997). White Lupin (Lupinus albus L.). Field Crops Res. 53, 147–160. doi:10.1016/S0378-4290(97)00028-2

Hyvönen, T., and Boyd, J. (1977). Origin, Ethnobotany and Agricultural Potential of the Winged Bean: Psophocarpus Tetragonolobus. Econ. Bot. 31, 180–188. doi:10.1007/BF02866589

Iizuka, N. (1990). Studies on Virus Diseases of Azuki Bean (Vigna angularis Wight) in Japan. Bull. Tohoku Nat. Agric. Exp. Station 82, 77–113.

Iqbal, M. M., Erkine, W., Berger, J. D., and Nelson, M. N. (2020). Phenotypic Characterisation and Linkage Mapping of Domestication Syndrome Traits in Yellow Lupin (Lupinus luteus L.). Theor. Appl. Genet. 133, 2975–2987. doi:10.1007/s00122-020-04365-9

Iqbal, M. M., Huynh, M., Uddal, A. J., Kilian, A., Dluhoš, P., Pátková, L., Nedlík, J., and Řehůřek, J. (2017). Construction of a Genetic Linkage Map and Genetic Analysis of the Domesticated Related Traits in rice bean (Vigna umbellata). Ann. Bot. 106, 927–944. doi:10.1093/aob/mcq188

Isemura, T., Kaga, A., Tomooka, N., Shimizu, T., and Vaughan, D. A. (2010). Construction of a Genetic Linkage Map and Genetic Analysis of the Domesticated Related Traits in Mungbean (Vigna radiata). PLoS One 7, e14304. doi:10.1371/journal.pone.0014304

Isobe, S., Kolliker, R., Hisano, H., Sasamoto, S., Wada, T., Klimenko, L., et al. (2009). Construction of a Consensus Linkage Map for Red Clover (Trifolium pratense L.). BMC Plant Biol. 9 (1), 1–11.

Istvánéves, J., Dhulsović, J., Dluhoš, P., Pátiková, L., Nedlík, J., and Řehůřková, J. (2017). Genome Classification and Mining of Molecular Markers Useful in Red clover (Trifolium pratense) Breeding. Front. Plant. Sci. 8, 367. doi:10.3389/fpls.2017.00367

Istvánéves, J., Jarolí, M., Křeneč, A., and Řehůrková, J. (2014). Genome Assembly and Annotation for Red clover (Trifolium Pratense; Fabaceae). Am. J. Bot. 101, 327–337.

Jasrotia, R. S., Yadav, P. K., Iqbal, M. A., Bhatt, S. B., Arora, V., Angadi, U. B., et al. (2019). VIGSatDB: Genome-wide Microsatellite DNA Marker Database of Three Species of Vigna for Germplasm Characterization and Improvement. Database. 2019 Jan 1. doi:10.1093/database/baz155

Jegadeesan, S., Raizada, A., Dhanasekar, P., and Suprasanna, P. (2021). Draft Genome Sequence of the Pulse Crop Blackgram [Vigna mungo (L.) Hepper] Reveals Potential R-Genes. Sci. Reps. 11, 1–0. doi:10.1038/s41598-021-06883-9

Jha, U. C. (2018). Modern Diseases and Insect–Pests of Rice Bean,” in Ricebean (Singapore: Springer), 347–353. doi:10.1007/978-981-15-5293-9_18

Kaur, T., Eswaran, M., Thangaraj, A., Meyyazhagan, A., Nehra, M., Raman, N. M., et al. (2019). Rice Bean (Vigna umbellata) Draft Genome Sequence: Unraveling the Late Flowering and Unpalatability Related Genomic Resources for Efficient Domestication of This Underutilized Crop. bioRxiv 2019 Jan 1. doi:10.1101/816595

Kaur, P., Bayer, P. E., Milec, Z., Vrana, J., Yuan, Y., Appels, R., et al. (2017). An Advanced Reference Genome of Trifolium Subterraneum L. Reveals Genes Related to Agronomic Performance. Plant Biotechnol. J. 15, 1034–1046. doi:10.1111/pbi.12697

Keatinge, J. D., Easdown, W. J., Yang, R. Y., Chadha, M. L., and Kim, T. S., Raveendar, S., Suresh, S., Lee, G. A., Lee, J. R., Cho, J. H., et al. (2015). Draft Genome Sequence of Adzuki Bean, Vigna Angularis. Sci. Reps. 5, 1–8. doi:10.1038/srep09069

Kasprzak, A., Sašťák, J., Janda, J., Doležel, J., Wolkó, B., and Naganowska, B. (2006). The Bacterial Artificial Chromosome (BAC) Library of the Narrow-Leafed Lupin (Lupinus angustifolius L.). Cell. Mol. Biol. Lett. 11, 396–407. doi:10.2478/s11658-006-0033-3

Katoch, R. (2020). “Common Diseases and Insect–Pests of Rice Bean,” in Ricebean (Singapore: Springer), 347–353. doi:10.1007/978-981-15-5293-9_18

Kim, T. S., Raveendar, S., Suresh, S., Lee, G. A., Lee, J. R., Cho, J. H., et al. (2015). Transcriptome Analysis of Two Vicia Sativa Subspecies: Mining Molecular Markers to Enhance Genomic Resources for Vetch Improvement. Genes 6 (4), 1164–1182. doi:10.3390/genes6041164

Kislev, M. E. (1989). Origins of the Cultivation of Lathyrus Sativus and L. Cicera (Fabaceae). Econ. Bot. 43, 262–270. doi:10.1007/BF02859868

Kitano-Okada, T., Ito, A., Koide, A., Nakamura, Y., Han, K. H., Shimada, K., et al. (2012). Anti-obesity Role of Adzuki Bean Extract Containing Polyphenols: In Vivo and In Vitro Effects. J. Sci. Food Agric. 92 (13), 2644–2651. doi:10.1002/jsfa.5680

Konduri, V., Godwin, I. D., and Liu, C. J. (2000). Genetic Mapping of the Ustilago Maydis Gene(s) in This Species. Theor. Appl. Genet. 100, 866–871. doi:10.1007/s001220051363

Korte, A., and Farlow, A. (2013). The Advantages and Limitations of Trait Analysis with GWAS: a Review. Plant Methods 9, 29. doi:10.1186/1746-4811-9-29
with Resistance to Bruchids (Callosobruchus spp.) in Wild Mungbean (Vigna radiata Var. sublobata) and Cultivated V. radiata through Genotyping by Sequencing and Quantitative Trait Locus Analysis. BMC Plant Biol. 16, 1–5.

Shapert, F. M., Cross, M., Ablett, G., Malory, S., Chivers, I. H., King, G. J., et al. (2013). High-throughput Sequencing and Mutagenesis to Accelerate the Domestication of Micronutrients in a New Food Crop. PLoS One 8, e82641. doi:10.1371/journal.pone.0082641

Sharma, S., Tyagi, A., Srivastava, H., Ramakrishna, G., Sharma, P., Sevanthi, A., M., et al. (2021). Exploring the Edible Gum (Galactomannan) Biosynthesis and its Regulation during Pod Developmental Stages in Cluster Bean Using Comparative Transcriptomic Approach. Sci. Reps. 11, 1–2. doi:10.1038/s41598-021-83507-3

Shegrov, A., Van Rensburg, W. J., and Adebola, P. O. (2013). Assessment of Genetic Variability in Bambara Groundnut (Vigna subterrenea L. Verdc.) Using Morphological Quantitative Traits. Academia J. Agric. Res. 1 (3), 045–051.

Sherrman, R. M., and Salzberg, S. L. (2020). Pan-genomics in the Human Genome Era. Nat. Rev. Genet. 21, 243–254. doi:10.1038/s41576-020-0210-7

Shi, R., Jiao, W., Yun, L., Zhang, Z., Zhang, X., Wang, Q., et al. (2021). Utilization of Transcriptome, Small RNA, and Degradome Sequencing to Provide Insights into Drought Stress and Water Restoring Treatment in Medicago Ruthenica. Front. Plant Sci. 12, 675903. doi:10.3389/fpls.2021.675903

Shirasawa, K., Kosugi, S., Sasaki, K., Ghelfi, A., Okazaki, K., Toyoda, A., et al. (2021b). Genome Features of Common Vetch (Vicia sativa) in Natural Habitats. bioRxiv. doi:10.1101/2021.03.09.434686

Shirasawa, K., Chahota, R., Hirakawa, H., Nagano, S., Nagasaki, H., Sharma, T. R., et al. (2021a). A Chromosome-Scale Draft Genome Sequence of Horsegram (Macrotyloma uniflorum). bioRxiv. doi:10.1101/2021.01.18.427074

Singh, T., Ramakrishnan, S., Mahanta, S. K., Tyagi, V. C., and Roy, A. K. (2018). Development of an SNP-Based High-Density Linkage Map and QTL Analysis. Theor. Appl. Genet. 136, 398. doi:10.1007/s00122-018-2768-2

Singh, V., G, A., Malory, S., Chivers, I. H., King, G. J., et al. (2013). High-throughput Sequencing and Mutagenesis to Accelerate the Domestication of Micronutrients in a New Food Crop. PLoS One 8, e82641. doi:10.1371/journal.pone.0082641

Tanwar, U. K., Pruthi, V., and Randhawa, G. S. (2017). RNA-seq of Guar (Cyamopsis tetragonolobus) in Natural Habitats. BMC Res. Notes 10, 1607. doi:10.1186/s12864-017-4334-2

Taylor, C. M., Kamphuis, L. G., Zhang, W., Garg, G., Berger, J. D., Mousavi-Darazmalleh, M., et al. (2019). INDEL Variation in the Regulatory Region of the Major Flowering Time Gene LanF1c Is Associated with Vernalization Response and Flowering Time in Narrow-leaved Lupin (Lupinus angustifolius L.). Plant Cell Environ. 42, 174–187. doi:10.1111/pce.13320

Taylor, C. M., Kamphuis, L. G., Zhang, W., Garg, G., Berger, J. D., Mousavi-Darazmalleh, M., et al. (2019). INDEL Variation in the Regulatory Region of the Major Flowering Time Gene LanF1c Is Associated with Vernalization Response and Flowering Time in Narrow-leaved Lupin (Lupinus angustifolius L.). Plant Cell Environ. 42, 174–187. doi:10.1111/pce.13320

Thirumaran, A. S., and Kanchana, S. (2000). “Role of Pulses in Human Diets,” in Pulses Production Strategies in Tamil Nadu (Coimbatore, India: Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University), 129. Available at: http://www.tnau.ac.in/tech/pulsebook.pdf (Accessed July 13, 2012).
Yin, L., Zhang, M., Wu, R., Chen, X., Liu, F., and Xing, B. (2021). Genome-wide Analysis of OSCA Gene Family Members in Vigna Radiata and Their Involvement in the Osmotic Response. *BMC Plant Biol.* 21, 1–2. doi:10.1186/s12870-021-03184-2

Yin, M., Zhang, S., Du, X., Mateo, R. G., Guo, W., Li, A., et al. (2021). Genomic Analysis of Medicago Rutherinica Provides Insights into its Tolerance to Abiotic Stress and Demographic History. *Mol. Ecol. Resour.* 21 (5), 1641–1657. doi:10.1111/1755-0998.13363

Yin, X., Yi, K., Zhao, Y., Hu, Y., Li, X., He, T., et al. (2020). Revealing the Full-Length Transcriptome of Caucasian clover Rhizome Development. *BMC Plant Biol.* 20, 1–5. doi:10.1186/s12870-020-02637-4

Yoshida, Y., Murubodee, R., Ogiso-Tanaka, E., Iseki, K., Isemura, T., Takahashi, Y., et al. (2016). Salt tolerance in wild relatives of adzuki bean, Vigna angularis (Willd.) Ohwi et Ohashi. *Genet. Resour. Crop Evol.* 63, 627–637. doi:10.1007/s10722-015-0272-0

Yoshida, J., Tomooka, N., Yee Khaing, T., Shantha, P. S., Naito, H., Matsuda, Y., et al. (2020). Unique Responses of Three Highly Salt-Tolerant Wild Vigna Species against Salt Stress. *Plant Prod. Sci.* 23, 114–128. doi:10.1080/1343943X.2019.1698968

You, M. P., Nichols, P. G., Katsumine, R., and Barbetti, M. J. (2021). Novel Disease Host Resistances in the World Core Collection of Trifolium Subterraneum. *Plant Dis.*. PDIS-09. doi:10.1094/pdis-09-20-1985-re

Yuan, J., Wang, B., and Wu, T. L. (2011). Quantitative Trait Loci (QTL) Mapping for Inflorescence Length Traits in *Lupinus purpureus* (L.) Sweet. *Afr. J. Biotechnol.* 10, 3558–3566.

Yundaeng, C., Somta, P., Alikul, K., Kongijumun, A., Kaga, A., and Tomooka, N. (2019). Construction of Genetic Linkage Map and Genome Dissection of Domestication-Related Traits of Moth Bean (*Vigna aconitifolia*), a Legume Crop of Arid Areas. *Mol. Genet. Genom.* 294, 621–635. doi:10.1007/s00438-019-01536-0

Zhang, F., and Batley, J. (2020). Exploring the Application of Wild Species for Crop Improvement in a Changing Climate. *Curr. Opin. Plant Biol.* 56, 218–222. doi:10.1016/j.pbi.2019.12.013

Zhang, H., Li, Y., and Zhu, J. K. (2018). Developing Naturally Stress-Resistant Crops for as Sustainable Agriculture. *Nat. Plants* 4, 989–996. doi:10.1038/s41477-018-0309-4

Zhang, H., Tian, H., Chen, M., Xiong, J., Cai, H., and Liu, Y. (2018). Transcriptome Analysis Reveals Potential Genes Involved in Flower Pigmentation in a Red-Flowered Mutant of white clover (*Trifolium repens* L.). *Genomics* 110, 191–200. doi:10.1016/j.ygeno.2017.09.011

Zhang, Q., Yan, Q., Yuan, X., Lin, Y., Chen, J., Wu, R., et al. (2021). Two Polygalacturonase-Inhibiting Proteins (VrPGIP) of Vigna Radiata Confer Resistance to Bruchids (*Callosobruchus* Spp). *J. Plant Physiol.* 258, 153376. doi:10.1016/j.jplph.2021.153376

Zhang, Y., Massel, K., Godwin, I. D., and Gao, C. (2018). Applications and Potential of Genome Editing in Crop Improvement. *Genome Biol.* 19, 1. doi:10.1186/s13059-018-1586-y

Zhou, G., Jian, J., Wang, P., Li, C., Tao, Y., Li, X., et al. (2018). Construction of an Ultra-high-density Consensus Genetic Map, and Enhancement of the Physical Map from Genome Sequencing in Lupinus angustifolius. *Theor. Appl. Genet.* 131, 209–223. doi:10.1007/s00122-017-2997-y

Zhou, G., Yang, H., Renshaw, D., Zou, M., Thomas, G., and Li, C. (2021). Identification of Gray Leaf Spot Disease Candidate Gene in Narrow-Leafed Lupin (*Lupinus angustifolius* L.). *Front. Genet.* 2021, 1375. doi:10.3389/fgene.2021.695791

Zhu, Z., Chen, H., Xie, K., Liu, C., Li, L., Liu, L., et al. (2020). Characterization of Drought-Responsive Transcriptome during Seed Germination in Adzuki Bean (*Vigna angularis* L.) by PacBio SMRT and Illumina Sequencing. *Front. Genet.* 11, 996. doi:10.3389/fgene.2020.00996

Zongo, E., Neya, B. J., Traoré, V. S. E., Palanga, E., Zabré, J., Barro, N., et al. (2018). Impact of Cowpea Mottle Virus on the Growth and Yield of Barnyard Groundnut (*Vigna subterranuana* (L.) Verdc.). *Amer. J. Plant Sci.* 9, 2053. doi:10.4246/ajps.2018.910149

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

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

Copyright © 2022 Jha, Nayyar, Parida, Bakır, von Wettberg and Siddique. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.