agriculture

MDPI

Review

The Exceptionally Large Genomes of the Fabeae Tribe: Comparative Genomics and Applications in Abiotic and Biotic Stress Studies

Carmen Santos *,† and Susana Trindade Leitão †

Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal; sleitao@itqb.unl.pt

* Correspondence: css@itqb.unl.pt

† These authors contributed equally to this work.

Abstract: The Fabeae tribe comprises five legume genera, which include some of the most ancient and important crops, like peas, lentils, and faba beans. Biotic and environmental stresses are major threats to the stable and high productivity of Fabeae crops. The use of omics resources can provide breeders with the tools needed to develop new crop varieties in a more efficient and sustainable way. However, the genomic efforts on Fabeae crops have lagged behind compared to other legume species, mainly due to their large genome size and repeat content. The first annotated chromosome-level reference genome assembly in Fabeae was published for pea (*Pisum sativum* cv. Caméor) in 2019. Since then, many efforts have been made to sequence the genome of other species from this tribe. Currently, 17 genomes of Fabeae species are available for the scientific community; five of them are at the chromosome level. Fundamental knowledge and molecular tools for breeding have been boosted on the legume resistance/tolerance against biotic and abiotic stresses by the availability of some of these recent reference genomes, especially the pea cv. Caméor genome. This review provides a comparison of the Fabeae tribe genomes available and an overview of recent accomplishments in their application in abiotic and biotic stress research.

Keywords: Fabeae tribe; genome sequencing; legumes; stress resistance; genomic resources

Citation: Santos, C.; Leitão, S.T. The Exceptionally Large Genomes of the Fabeae Tribe: Comparative Genomics and Applications in Abiotic and Biotic Stress Studies. *Agriculture* **2024**, *14*, 77. https://doi.org/10.3390/agriculture14010077

Academic Editor: Rodomiro Ortiz

Received: 15 November 2023
Revised: 22 December 2023
Accepted: 28 December 2023
Published: 30 December 2023

1. Introduction

The legume family (Fabaceae) includes some of the most important crops for human and animal nutrition worldwide and contributes to environmentally friendly agriculture due to the symbiotic natural ability to fix atmospheric nitrogen. The current focus in legume research is on accelerating genetic gains related to yield, nutritional quality, and stress resistance/tolerance [1].

Legume cultivation is affected by lower crop yields due to biotic and abiotic stress factors [2,3]. Plant response to these stresses is known to involve several genes [4]. Genome sequencing allows for the decoding and understanding of the mechanism or genetic basis for the functional characterization of genes that have a great impact on crop improvement. The development of next-generation sequencing (NGS) technologies and bioinformatics brings the possibility of sequencing any genome or transcriptome [5]. Among legumes, the genome sequence of the soybean (*Glycine max*) was the first to be completed in 2010 [6]. Other model, crop, and orphan legume species have the whole-genome sequences available since the last decade, including barrel medic (*Medicago truncatula*) [7], common bean (*Phaseolus vulgaris*) [8], mung bean (*Vigna radiata*) [9], and pigeonpea (*Cajanus cajan*) [10]. Only recently, larger-sized genomes of important crop legumes such as faba bean (*Vicia faba*) [11], pea (*Pisum sativum*) [12–14], and lentils (*Lens culinaris*) [15] have been reported.

The availability of genome sequences and large-scale genomic resources facilitates the advances in functional genomics and molecular breeding, which have enhanced the efficiency and resolution of genetic mapping, marker-trait association, and genome-editing studies in legume research [16–20]. Nevertheless, genome assemblies developed for several legume crops are based on short reads (Illumina sequencing-based), have uneven sequencing depths and are fragmented [21]. Whole-genome duplication (WGD) events and extensive repeat content are the main challenges for assembling legume plant genomes.

The majority of legume genome crop genome sizes range from 0.3 to 1.5 Gbp [22]. In the Fabaceae family, the tribe Fabeae presents the largest amounts of DNA within a haploid chromosome set C-values, varying between 1.8 to 13 Gbp [11], exceptionally larger genomes compared to the closely related Trifolieae (genome size ~1.05 Gb) and Cicereae (genome size ~1.27 Gb) [23]. This variation in size and expansion/contraction has occurred during its evolution due to the presence of repetitive DNA. The tribe Fabeae comprises five genera (*Lathyrus* L., *Lens* Mill., *Vicia* L., *Pisum* L., and *Vavilovia* Fed.) and includes several important crops, such as peas, lentils, and faba beans [24]. In addition to crops of worldwide importance, the tribe also contains minor crop species (e.g., *Lathyrus sativus* L.), forage or green manure species (e.g., *Vicia sativa* L. and *V. villosa* L.), and ornamental species (e.g., *L. odoratus* L. and *L. tuberosus*). Despite the importance of this group, only recently, the genome sequence of some species began to emerge and became available to the scientific community: *Pisum sativum* [12–14], *Lathyrus sativus* [25,26], *Lathyrus tuberosus* [27], *Vicia sativa* [28–30], *Vicia faba* [11], *Lens culinaris* and *Lens ervoides* [15]. This review aims to compare the genomes available from the Fabae tribe and explore their application in abiotic and biotic stress studies.

2. Fabeae Tribe Genomes

2.1. Pisum sativum

Pea, commonly known as dry, green, or field pea, is the second most important grain legume in the world after the common bean, being a valuable source of dietary proteins, mineral nutrients, complex starch, and fibers with proven health benefits [31]. Pea was the original model organism used in Mendel's discovery of the inheritance laws, making it the foundation of modern plant genetics [32]. After Mendel studies, subsequent pea genomics was neglected compared with other plant species. Recently, in 2019, the genome of *Pisum sativum* cv. Caméor was published as the first annotated chromosome-level reference genome assembly in the tribe Fabeae [13]. Since then, many efforts have been made to sequence the genome of other species in this tribe.

Regarding the *Pisum* genus, the National Center for Biotechnology Information (NCBI) repository reports six genomes, all for *P. sativum* (Table 1): two genomes at the chromosome level [12,13], two genomes at the scaffold level (unpublished) and two at the contig level [14,33]. The two chromosome-level reference genome assemblies include the "Pisum_sativum_v1a," of a French pea cultivar "Caméor" [13], and the more recent reference genome for pea "CAAS_Psat_ZW6_1.0" of the Chinese pea variety Zhongwan 6 (ZW6) [12]. The other pea contig level genome assemblies comprise the "PSA_r1.0" of the pea cultivar JI128, provided by John Innes Centre, , UK [14], the "ASM301357v1" of a pea cultivar "Gradus No 2" from the Earlham InstituteUK, and the "ASM2453087v1" of a Russian pea cultivar "Frisson" (Table 1). The NCBI repository also includes the *Pisum sativum* CEN6 assembly (GenBank GCA_947076115.1), but this assembly only covered a 177.6 Mbp region of pea chromosome 6 and not the complete genome [33].

Table 1. Summary of characteristics and statistics of the Fabeae genome assemblies. n.a. means not available.

Genome	Online Platform or Genome Database	%GC Content	Biological Samples	Size Genome (Gbp)	Ploidy	Genomic DNA Sequencing Technology	Genome Coverage (Fold)	Assembled Scaffolds (nr/Length Mbp)	Contigs (nr/Length Mbp)	N50 Contigs (kbp)	Complete BUSCO Assembly (%)	Protein-Coding Gene Models (nr/Average Length Kbp)	Complete BUSCO Annotation (%)
Pisum sativum Caméor	https://urgi.versailles.inra.fr/Species/Pisum/Pea-Genome-project, accessed on 15 November 2023	37.6	French pea cultivar Caméor	4.45	2n = 14	Illumina and PacBio RSII	281	24,623/ 3919.1	218,010/ 3159.58	37.931	96.78	44,756/124.6	93.99
Pisum sativum ZW6	https://www.peagdb.com/index/, accessed on 15 November 2023	n.a.	Chinese pea cultivar Zhongwan 6 (ZW6)	4.28	2n = 14	Illumina HiSeqX Tem, SMRT- PacBio seq, Illumina Nova Seq 6000	85.2	1579/ 3796.7	2402/ 3786.4	8980	99.38	47,526/121.84	97.77
Pisum sativum JI128 PSA_r1.0	PRJDB10540 (NCBI), accessed on 15 November 2023	36.5	JI128 (Norwich, United Kingdom)	4.5	2n = 14	MinION (Nanopore)	44	n.a.	117,981/ 297.04	51.222	95.4	531,242	n.a.
Pisum sativum Gradus	PRJNA432052 (NCBI), accessed on 15 November 2023	37.5	SAMN08429545 Gradus No 2 (Earlham Institute, United Kingdom)	4.3	2n = 14	Illumina HiSeq	86	5,449,423/ n.a.	5,465,676/ 4274.31	460.1	91.5	n.a.	n.a.
Pisum sativum Frisson	PRJNA812957 and PRJNA853105 (NCBI), accessed on 15 November 2023	38	ASM2453087v1 Russian pea cultivar Frisson	3.7	2n = 14	Oxford Nanopore MinION; Illumina HiSeq	100	2254/ n.a.	2355/ n.a.	5000	n.a.	n.a.	n.a.
Lathyrus sativus Eiv1	PRJEB33571 (NCBI), accessed on 15 November 2023	38.3	European cultivar LS007	8.12 (including 1.1 Gbp of Ns)	2n = 14	Illumina HiSeq	n.a.	669,893 (>1 kbp)		59.728	82.8 (Fabales)	33,819/ 1.454	79.9 (Fabales)
Lathyrus sativus Rbp	https://zenodo.org/record/7390878, accessed on 15 November 2023	38.8	European cultivar LS007	6.52	2n = 14	Oxford Nanopore Technologies long-read	45.44	162,994 (>1 kbp)		157.998	89.6 (Fabales)	63,922	82.6 (Fabales)
Lathyrus sativus Pusa-24	https://lathyrus genome.nabi.res.in/, accessed on 15 November 2023	38.32	Indican cultivar Pusa-24	6.62	2n = 15	Illumina HiSeq. 2500 platform; PacBio Sequel (I) platform	n.a.	25,411/ 3805	80,744/ 3800	78.3	95.96 (Fabales)	50,106	n.a.
Lathyrus tuberosus	PRJNA810344 (NCBI), accessed on 15 November 2023	38.3	*Lathyrus tuberosus* NL20	6.8	2n = 14	PacBio HiFi sequencing	30	n.a.	1353/ 6791	11,100	96.6 (Fabales)	n.a.	n.a.
Lens culinaris Lcu.2RBY	https://knowpulse.usask.ca/genome-assembly/Lcu.2RBY, accessed on 15 November 2023	n.a.	Canadian red cultivar CDC Redberry	3.92	2n = 14	PacBio SMRT + Oxford Nanopore Technology	34 + 20	3760	6094	1300	94.4	39,778	94
Lens ervoides Ler.1DRT	n.a.	n.a.	Wild accession IG 72815	2.9	2n = 14	Oxford Nanopore Technology	52	2870	2291	4700	95.1	37,045	95
Vicia faba	https://projects.au.dk/fabagenome//, accessed on 15 November 2023	n.a.	Hedin/2	13	2n = 12	Hedin/2: PacBio HiFi long reads	20	3986/ 11,900	10,721	2700	96.3	34,221	93.3

Vicia faba	https://projects.au.dk/fabagenome//, accessed on 15 November 2023	n.a.	German cultivar Tiffany	13	2n = 12	Tiffany: PacBio HiFi reads using KAT (v2.4.2)	10	1971/11,400	14,378	1600	91.6	34,043	93
Vicia sativa KSR5 VSA_r1.0	https://plantgarden.jp/ja/list/t3908/genome/t3908.G001, accessed on 15 November 2023	n.a.	Japanese standard inbred line- KSR5	1.8	2n = 14	HiSeq2000 Illumina (paired-end) and RSII instrument (PacBio)	105	6,487,000/2500	54,083 (>1 kbp)	901	77.5	31,146	94.1
Vicia sativa Studenica ASM2176476v1	http://gigadb.org/dataset/view/id/100954/File_page/2, accessed on 15 November 2023	35.6	Australian cultivar Studenica (*V. sativa* subsp. *Sativa*)	1.65–1.77	2n = 12	PromethION, Illumina NovaSeq 6000 Sequencing System, and Illumina HiSeq X Tem	124.3 + 43.71 + 98.18	n.a.	9990	659	97.8	53,218	95.6
Vicia sativa	PRJNA730328 (NCBI), accessed on 15 November 2023	35.94	Chinese cultivar Lanjian No. 1	1.568	2n = 12	Illumina NovaSeq 6000 Sequencing System/PE150	n.a.	3,754,145/1,516,858,186	4,227,942	1.245	n.a.	n.a.	n.a.
Vicia villosa Vvill1.0	PRJNA868110 PRJNA833581 (NCBI), accessed on 15 November 2023	35.62	HV-30 AU merit cultivar	1.629	2n = 14	Illumina short-read NextSeq 500 WGS and Hi-C, PacBio HiFi sequencing Sequel II platform	74 + 21 + 43	1888	5373/1384	604	99	53,321	81.8

Pisum sativum (2n = 2x = 14) has a large and complex genome, with an estimated size ranging between 3.7 and 4.5 Gbp, containing about 85% repetitive DNA. Sequencing coverage in pea genomes has varied from 44 (cv. JI128) to 281-fold (cv. Caméor) (Table 1). To overcome the challenges of size and repetitiveness, a combination of long-read sequencing developed by Pacific Biosciences or Oxford Nanopore with Illumina short-read DNA data has been useful for genome assembly[12,13,34]. Since long-read sequencing has a higher error rate than short-read Illumina or Ion sequencing, hybrid sequencing or increasing coverage can greatly improve the accuracy of genome assembly.

Assembly complementary approaches have been combined to obtain the pea reference genomes, including high-throughput assembly analyses and manual curation steps [12,13]. In the pea cv. Caméor genome, Illumina short-read sequences were assembled into contigs and then combined into scaffolds using long-range PacBio RSII (Pacific Biosciences) sequences and whole-genome profiling of a bacterial artificial chromosome (BAC) library. Scaffolds were manually curated for inter and intrachromosomal chimeras using sequences obtained from single chromosomes isolated by flow cytometry and an ultra-high-density skim genetic map obtained by genotyping-by-sequencing. Curated scaffolds were then integrated into super-scaffolds using BioNano(San Diego, Califórnia, EUA) genomic maps (optical maps from BioNano Genomics). Finally, the seven pseudomolecules representing the seven pea chromosomes were obtained by anchoring super-scaffolds onto high-density genetic maps [13].

For the pea cv. ZW6 genome, a novel assembly was generated based on the PacBio SMRT long-read sequencing, combining 10x Genomics scaffolding, BioNano optical mapping, chromosome conformation capture (Hi-C) scaffolding, and Illumina technologies [12]. The initial assembly was polished by iterative scaffolding and manual curation, and the final assembly was anchored into seven chromosome-level pseudomolecules, two organelle genomes, and 1572 unplaced contigs. The pea cv. ZW6 assembly strategy resulted in an improved reference genome in continuity and quality of complex repeat regions and transposable elements that remained as gaps in the pea cv. Caméor genome [12,13]. The total size of anchored contigs was 3719.6 Mb, constituting 97.96% of the pea cv. ZW6 genome, whereas anchored contigs constituted only 82.51% of the pea cv. Caméor assembly.

The cumulative length of unknown sequences was 10.3 Mb for pea cv. ZW6, which was smaller than the 760.8 Mb observed in pea cv. Caméor.

Less complex assembly methods were applied for the genome assembly of pea cv. JI128, where only the long reads were used. Sequencing errors in the assembly were then corrected, for which Illumina and MGI short reads were employed. Assembly was then anchored to the reference genome of cv. Caméor and Gradus No2 assembly using SNPs obtained by ddRAD (double-digest restriction site-associated DNA)-Seq, as anchors. Information about the methodology for genome assembly of the Gradus No2 and Frisson is not available yet (unpublished studies).

Annotation of pea genomes has been focused on the identification of repetitive elements and prediction of coding genes. The total length of repetitive elements in pea cv. ZW6 was 3249.5 Mbp, larger than in pea cv. JI128 (2825 Mbp) and in pea cv. Caméor (2662.5 Mbp). The majority of the repetitive elements found in these genomes corresponded to transposable elements (TE), from which the long terminal repeats (LTR) retrotransposons were the most abundant in the three annotated pea genomes. In particular, Ogre elements are a type of Ty3/Gypsy LTR retrotransposon first discovered in legumes [35,36] characterized by their large size (reaching 25 kbp) and the presence of an additional ORF (open reading frame) encoding a protein of unknown function located upstream of the gag gene, usually present in LTR retrotransposons [37]. In the Fabeae tribe, Ogre elements typically make up about 40% of the entire genome (22.5 − 64.7%), and they are considered the major force driving the evolution of the genome size in this tribe [38]. This high level of repetition emphasizes the difficulty of assembling the genomes of Fabeae members and the need for long sequence reads that can span repetitive regions. In pea cv. Caméor, LTRs represented 72.7% of the genome (1,707,747 copies), with Ogre elements being the major representative [13].

The pea cv. ZW6 genome showed a much higher number of full-length LTR (62,978) than pea cv. Caméor (9999) and a higher percentage of active and longer LTRs [12]. The difference in the number of these repetitive elements identified in both pea genomes may justify the differences in genome assembly gap size between the pea cv. ZW6 6 (10.3 Mbp) and pea cv. Caméor (760.8 Mbp). Indeed, the LTR assembly index (LAI), which evaluates assembly continuity, indicated a substantial improvement in LTR completeness for pea cv.ZW6 (LAI = 13.31) compared to pea cv. Caméor (LAI = 2.09). In pea cv. JI128 assembly, LTRs occupied about 69% (2284 Mbp) of the genome assembly [14].

Pea protein-coding genes have been annotated using a combination of ab initio, homology-based, and transcriptome-based prediction. In the pea cv. In the JI128 genome, 531,242 potential protein-coding genes were predicted, and the pea genome had more predicted genes [14]. Similar numbers of predicted coding genes were reported for pea cv. Caméor (44,756 complete and 29 truncated genes) and pea cv. ZW6 (47,526). Also, similar average gene lengths were obtained between these two pea genomes: 124.6 Mbp for pea cv. Caméor, and 121.84 Mbp for pea cv. ZW6, respectively.

Pea genome assemblies and gene set completeness have been assessed by Benchmarking Universal Single-Copy Orthologs (BUSCO) using the "embryophyta_odb10" model. Assembly completeness varied from 91.5 (Gradus no. 2) to 99.38% for pea cv. ZW6 (Table 1). The protein mode BUSCO completeness of annotated genes was also higher in pea cv. ZW6 (97.77%) than that in pea cv. Caméor (93.99%) [12].

High-quality reference genomes and annotations provide fundamental resources for functional genomics, characterizing genetic traits, and breeding in crops. The pea cv. Caméor and ZW6 genomes revealed insights into legume genome evolution and domestication and provided important resources for functional genomics and crop breeding [13]. Indeed, the availability of the first pea assembly genome (cv. Caméor) revealed genomic rearrangements across legumes and suggested a major role for repetitive elements in pea genome evolution [13]. Phylogenetics and paleogenomics showed genomic rearrangements across legumes and suggested a major role for repetitive elements in pea genome evolution. Compared to other sequenced legume genomes, the pea genome shows intense

gene dynamics, most likely associated with genome size expansion when the Fabeae diverged from its sister tribes [13]. The pea genome is differentiated from its Ancestral Galegoid Karyotype through at least three chromosomal fissions, four fusions, and a translocation between chromosomes Ps1 and Ps5 [13]. Genome expansion in plants is primarily driven by WGD events and the proliferation of TEs. LTRs account for most of the genome size differences between *P. sativum* and *M. truncatula*, *Trifolium pratense* L., *L. japonicus*, *P. vulgaris*, or *G. max* [13].

Besides the pea cv. ZW6 reference genome, Yang et al. [39] also reported a pea pan-genome based on 118 accessions. Pan-genomes comprise the genomic diversity of a species, including core, variable, and unique genes found in a large set of individuals. The whole-genome resequencing and transcriptome data obtained for JI128, JI4, and an F2 population (JI4 × JI128, $n = 167$) may facilitate pan-genomic future studies in pea [14].

2.2. Lathyrus spp. (L. sativus and L. tuberosus)

After pea cv. Caméor reference genome publication, Emmrich et al. [40] reported a draft genome of a close phylogenetic relative, *Lathyrus sativus* (grass pea), using a European accession (LS007). Indeed, these authors generated two assemblies of this genome: the EIv1 using short-read Illumina paired-end sequencing and the Rbp combining the Illumina paired-end with long-read Oxford Nanopore Technologies sequencing (PromethION nanopore). More recently, the Rbp genome assembly was also annotated [26]. The Rbp genome assembly and annotation (at scaffold level) is deposited in NCBI as "LS007 EIv1 GenomeAssembly" and deposited on Zenodo [https://zenodo.org/record/7390878], accessed on 15 November 2023. Another grass pea genome assembly became available at the scaffold level, the "ASM2687324v1" of an Indian grass pea cultivar (Pusa-24) [25]. The Pusa-24 *L. sativus* genome is also available at https://lathyrusgenome.nabi.res.in/index.html online platform, accessed on 15 November 2023, which provides BLAST tools, an ortholog search tool, and direct download of the assembled genome, and associated files. Additionally, an assembly at the contig level is also available for *L. tuberosus* [27].

Grass pea ($2n = 2x = 14$) is an important food security crop due to its resilience under hostile environmental conditions, such as drought [41,42]. The crop has a long history in agriculture, but its cultivation has declined, probably due to the presence of the neurotoxin β-ODAP [41,43]. This neurotoxin has been connected with a neurological disease (neurolathyrism) that occurs when grass pea is the main source of nutrition in an unbalanced diet for prolonged periods [44]. The recent available genetic and genomic resources may facilitate breeding for β-ODAP low content. Still, resistance to biotic stress is the main target for breeding in *Lathyrus* spp. [43,45].

Another *Lathyrus* species, *L. tuberosus* ("tuber vetchling", $2n = 2x = 14$), has been cultivated for flavorings, fodder (from leaves), oil (from seeds), and as a potato-like food (from tubers) [46]. To aid future domestication programs for this legume plant and facilitate evolutionary studies of tuber formation, genetic and genomic resources are needed for this tuberous legume [27].

The genome size of *Lathyrus* species varies from approximately 3.5 to 14.3 Gbp (*Lathyrus vestitus*), the largest known diploid legume genome [22]. From the genome assemblies available, *L. sativus* genome size varied from 6.517 ± 0.023 Gbp to 6.620 ± 0.095 Gbp, using the pea genome as standard [25,26]. The length of the EIv1 assembly is 8.12 Gbp, including 1.9 Gbp of Ns in scaffolds [40]. For *L. tuberosus*, the estimated genome size is 6.8 Gbp [27].

Hybrid assembly methodologies have been used in the *L. sativus* assembly genomes available, involving short sequencing data (Illumina) with long sequencing data (Pacific Biosciences, Menlo Park, Califórnia, EUA (PacBio) or Oxford Nanopore Technologies, Oxônia, Reino Unido (PromethION)) [25,26,40], while the PacBio was chosen to generate the *L. tuberosus* draft genome assembly [27]. For Rbp assembly, after filtering for reads >5 kbp and using a long-read assembler, Redbean produced an assembly of 6.2 Gbp based on 34.6X coverage. The resulting assembly contained 162,985 contigs, with a contig N50 of 155.574 kbp and a GC fraction of 38.8%. Following polishing with more assembly

algorithms, incorporating 49.7x coverage of paired-end Illumina HiSeq data resulted in an assembly of 6.237 Gbp in 162,994 contigs with an N50 of 157.998 kbp. Hybrid assembly of Pusa-24 was carried out using the default parameters in MaSuRCA [47]. The contig-level assembly covered 3.8 Gbp of the genome with a contig N50 value of 78.27 kbp and a GC fraction of 38.32% (Table 1).

Sequence reads from *L. tuberosus* were assembled with HIFIASM genome assembler [48], resulting in 1353 contigs with a total length of 6.8 Gbp, a contig N50 of 11.1 Mbp and a GC content of 38.3%, including five mitochondrial and a single chloroplast contigs [27]. Assembly BUSCO analysis showed 96.6% completeness (Fabales). Since *L. tuberosus* is thought to be an obligate outcrosser, heterozygous regions seem to be present in the assembly, which was indicated by the high degree of BUSCO groups (up to 30%) [27].

Due to the high level of fragmentation, Hi-C scaffolding of *L. sativus* Rbp assembly only resolved 42.7% of the assembly assigned to 7 chromosome-scale and 2 sub-chromosome-scale scaffolds [26]. The *L. sativus* cv. The Pusa-24 assembly genome was scaffolded by aligning Pusa-24 contigs to the *Pisum sativum* cv. Caméor v1a assembly. The scaffolded assembly contained seven chromosome-sized scaffolds and 25,404 contigs. The N50 value of the scaffolded assembly was 421.39 Mbp. This approach results in loss of cv. Pusa-24 genomic regions that were not mapped to the pea genome and may miss chromosomal rearrangements between pea cv. Caméor and grass pea cv. Pusa-24. In particular, it is relevant to clarify if the pea translocation between chromosomes Ps1 and Ps5 compared to the ancestral Galegoid karyotype [13], which is present in *Pisum sativum sativum* lineage, but not in *P. sativum elatius*, is also present in grass pea. Given the limitations in both *L. sativus* (Rbp and Pusa-24) assembly approaches and consequent limitations on annotation procedures, a new *L. sativus* genome assembly is still needed for future research on this promising species.

Annotation of *Lathyrus* genomes is only available for *L. sativus* and includes the identification of repetitive elements and prediction of coding genes. For Rbp genome assembly, annotation, and downstream analysis were conducted with the complete (6.2 Gbp assembly), unscaffolded assembly, aiming to not lose sequences of potentially important genes that were not scaffolded [26]. Gene models were predicted by an evidence-guided annotation approach incorporating RNA-seq of the three grass pea genotypes and cross-species protein alignments. In total, 30,167 high-confidence protein-coding genes and 15,307 low-confidence protein-coding genes were identified. For the Pusa-24 assembly genome, gene models were predicted by combining ab initio gene prediction, homology-based, and RNA-seq data. The final gene set consisted of 50,106 genes. The protein mode BUSCO completeness of annotated genes was higher in *L. sativus* Pusa-24 (96.0%) than in that Rbp assembly (82.6%), using Fabales as a model. Repetitive sequences represented 80.61% in Rbp assembly and 83.31% in Pusa-24 assembly. LTR elements were the most abundant in both grass pea assemblies, accounting for 78.3% of the contig lengths in Rbp assembly and 37.58% of the whole genome of Pusa-24 [25,26].

Further, to confirm the validity and quality of gene prediction of the *Lathyrus sativus* assemblies, researchers have searched for genes that contribute to β-ODAP biosynthesis. Since β-ODAP, an endogenous non-protein amino acid, is present exclusively in *Lathyrus* spp. and not in the other legumes, identification of the biosynthetic genes of this non-protein amino acid in the current genome assemblies further affirms the quality of the gene prediction. Edwards et al. [26] have elucidated the biosynthetic pathway leading to the β-ODAP formation. The final reaction of the pathway depends on an interaction between *L. sativus* acyl-activating enzyme 3 (AAE3) and a BAHD-acyltransferase (BOS) that forms a metabolon activated by CoA to produce β-L-ODAP. AAE3 participates in oxalate catabolism across various plant species, including *M. truncatula*. However, in *L. sativus*, an evolutionary repurposing of this established pathway results in the generation of β-L-ODAP. This alteration potentially offers an alternative method for eliminating oxalate [26]. Rajarammohan et al. [25] have identified the assembly of most of the known genes associated with the β-ODAP biosynthetic pathway in the Pusa-24 genome: viz. Serine O-

acetyltransferase (SAT), cysteine synthase (CS), cyanoalanine synthase (CAS), nitrilase, β-ODAP synthetase (BOS), oxalyl-CoA synthetase (OCS), and oxalate decarboxylase (ODC). Despite the assembly limitations discussed above, the results on the prediction of genes from the biosynthesis of β-ODAP are indicative of the high-quality overall completeness of the predicted gene set of the *L. sativus* genome assemblies and may be used for other gene discovery studies.

2.3. Vicia spp. (V. faba, V. sativa, and V. villosa)

The genus *Vicia* L. counts 150–210 species distributed across Europe, Asia, and North America, but mainly in the Mediterranean [49]. The best-known species of this genus is the faba bean or broad bean (*V. faba*, 2n = 2x = 12), an important food legume and fodder crop with nutritional relevance in the human diet. Faba bean is one of the earliest domesticated crops, dating back over 10,000 years in the near East. While no known wild progenitor exists today, it holds promise as a solution to future protein demands due to its high yield and protein content (around 29%), surpassing other cool-season pulses like peas, lentils, and chickpeas [50]. However, the faba bean has posed challenges for breeders because of its partially allogamous mating system, large genome size, and low seed multiplication rate.

Other vetches species, such as common vetch or spring vetch (*V. sativa*) and hairy vetch (*V. villosa*), with smaller protein-rich seeds, are normally used for animal feed (forage, fodder, pasture, silage, hay) and as cover crop due to their fast biomass production.

Four *Vicia* assembled genomes are available for the research community in NCBI: one of *V. faba* (cv. Hedin/2 genome v1), two of *V. sativa* (ASM2176476v1, and VSA_r1.0), and one of *V. villosa* (Vvil1.0). Their sizes range from 1.7 (*V. sativa*) to 11.9 Gbp (*V. faba*). *Vicia fava* Hedin/2 genome v1, *V. sativa* ASM2176476v1, and *V. villosa* Vvil1.0 are chromosome-level assemblies and the current reference genomes for the three species (Table 1). The assembly *V. sativa* VSA_r1.0 is a scaffold-level assembly. VfEP_Reference-Unigene is also listed in NCBI as a *V. fava* genome; however, it corresponds to a transcriptomic study. The raw sequence reads of *V. sativa* cv. Lanjian No. 1 is also included in NCBI in the project PRJNA730328. *Vicia faba* cv. Hedin/2 reference genome v1, the *V. sativa* ASM2176476v1, VSA_r1.0, and *V. villosa* Vvil1.0 are annotated. This genotype was chosen as a reference for the genome sequencing due to its high autofertility, productivity, early maturing spring habit, and homozygosity. Despite not being available in NCBI, there is a second faba bean genome assembly using the German variety, Tiffany [11]. The Faba Bean Genome Consortium made available the genome assemblies files and genome browser tools for both cv. Hedin/e and Tiffany at www.fabagenome.dk, accessed on 15 November 2023.

The Faba bean genome is massive, primarily due to the proliferation of repeat elements like retrotransposons and satellite repeats [11]. Despite its enormous size (11.9 Gbp for Hedin/2 and 11.4 Gbp for cv. Tiffany), it is surprisingly compact in terms of gene distribution, with genes and recombination events evenly dispersed across chromosomes. This compactness, however, is influenced by substantial copy number variation driven by tandem duplication. The faba bean genome is also noteworthy for being one of the largest diploid field crops, with its largest chromosome holding the equivalent of an entire human genome. Notably, there is no evidence of recent whole-genome duplication in the faba bean genome [11]. Nevertheless, more genes are duplicated in the faba bean genome compared to peas and lentils. Similar to other Fabeae members, LTR retrotransposons are an extensive part of the genome, almost 80% of the Hedin/2 and Tiffany assemblies [11]. LAI score of 10.5 was calculated for Hedin/2 assembly, providing evidence for its contiguity.

The genome sequence of cv. Hedin/2 was obtained using PacBio HiFi long reads, resulting in 20-fold coverage, N50 of 2700 kbp, a total contig number of 10,721, and 34,221 protein-coding genes (Table 1).

The GC content was 37.3%. The annotation was performed using RNA-seq data from nine tissues. Gene order was highly collinear and syntenic with *M. truncatula* and *P. sativum*. The gene model predictions captured 96% of single-copy conserved orthologues

according to the BUSCO analysis with the Embryophyta database. To further validate the gene annotation, several *M. truncatula* genes related to symbiosis with rhizobia or arbuscular mycorrhizal fungi were aligned, and putative orthologues were found for all those genes [11].

A similar number of gene models was also predicted in the assembly of cv. Tiffany using HiFi data (ten-fold coverage, N50 of 1600 kbp, a total contig number of 14,378, and 34,043 protein-coding genes, Table 1) [11]. Genetic mapping and Hi-C data were used to place a large portion of the genome (94%) into chromosomal pseudomolecules, providing a valuable frame of reference for genetic mapping, gene expression profiling, and comparative genomics. A higher proportion of BUSCO duplicate genes was obtained in Tiffany than in Hedin/2, leading to the need for the removal of ~10% of short-contig sequences in Tiffany to reach a similar low gene duplication level as in Hedin/2 (0.37%). For the Hedin/2 annotation, 69% of gene models were supported by RNA-seq data. In addition, 93.3% and 93% of gene models for Hedin/2 and Tiffany had similarities to proteins of close relatives (pea, lentil, Medicago, e-value cutoff 1×10^{-5}). Close to 80% of the Hedin/2 assembly was annotated as transposon-derived. Most of them (63.7%) are LTRs. Indeed, around 44% of the genome is made of *Gypsy* elements from the *Ogre* LTR retrotransposons family. The fact that the length of each Ogre element is up to 35 kbp partially explains the giant size of the faba bean genome [11].

Regarding *V. sativa*, three high-quality chromosome-level reference genomes were recently assembled using the cultivars Studenica, KSR5, and Lanjian No. 1 [28–30]. For this species, at least three different chromosome numbers (2n = 10, 12, and 14) are described, making 2n = 12 the best-characterized karyotype [29]. For Studenica *V. sativa*, Xi and colleagues [30] used a combination of long-read Oxford Nanopore sequencing, short-read Illumina sequencing, and high-throughput chromosome conformation data (CHiCAGO and Hi-C) analysis. The sequencing was performed on a PromethION and NovoSeq 6000 (Illumina NovaSeq 6000 Sequencing System). GenomeScope analysis estimated a genome-wide heterozygosity level of 0.09% and a genome size of 1.61 Gbp, slightly smaller than the flow cytometry estimate of 1.77 Gbp. The final genome assembly comprises six pseudo-chromosomes, with 1.65 Gbp of contigs anchored to them, leaving 10 contigs (overall 334,511 bp length) unassigned. The overall genome size was calculated as 1.653 Gbp, with a GC content of 35.6% (Table 1). The BUSCO analysis indicated a genome completeness of 98% of the dicotyledonous orthologs, with 9% representing duplicated genes. A total of 9990 assembled contigs were obtained with an N50 value of 685 kbp. RNA-seq analysis and gene modeling enabled the annotation of 53,218 protein-coding genes. BUSCO analysis identified 95.6% (Fabales) of completeness. The LTR elements constitute 64.2% (22.5% are *Ogre* elements) of this genome [30]. The high value of LAI (12.96) calculated suggests a reference quality of the Studenica *V. sativa* genome. Another draft genome sequence (VSA_r1.0) for common vetch was developed by Shirasawa et al. [29] using the Japanese standard inbred line KSR5, HiSeq2000 (Illumina), and RSII (PacBio) instruments. The genome size was estimated to be 1.769 Gbp, and the assembly size was 1.541 Gbp. The total number of scaffolds was 54,083, and the N50 was 90.1 kbp. Around 50% of the genome length is constituted by repeat sequences, 13% of which are assigned as LTR elements. According to BUSCO analysis, 77.5% of the genome was completely assembled, and the ortholog completion was 94.1%. In addition, Ma and colleagues [51] carried out a genome survey sequencing of common vetch on the Chinese cultivar Lanjian No. 1. The Illumina NovaSeq 6000 instrument with PE 150 sequencing methods was used to construct and sequence a genomic paired-end library with 300–400 bp short-inserts. A total of 79.84 Gbp of high-quality sequence data were obtained and assembled into 3,754,145 scaffolds, with an N50 length of 3556 bp. The K-mer analyses estimated the genome size to be 1.57 Gbp, the heterozygosity rate to be 0.4345, and the GC content to be 35% (Table 1). These results are very close to the ones obtained for the Studenica cultivar [28]. The assembled scaffolds were employed for genomic microsatellite (SSR) search, and 76,810 putative SSRs were identified.

Recently, Fuller et al. [52] disclosed a high-quality reference genome assembly for hairy vetch, *V. villosa*, using the cultivar AU Merit. They chose this cultivar because it can be clonally propagated in tissue culture. The genomic short-read libraries were sequenced on an Illumina NextSeq 500 instrument. The GenomeScope tool estimated the haploid genome size to be 0.883 Gbp, with an elevated heterozygosity of 3.14%, likely a result of the cross-fertilizing nature of *V. villosa*. Hi-C genome scaffolding was employed to create chromosome-scale scaffolds. The resulting Vvill1.0 assembly consisted of 1888 scaffolds totaling 2,034,988,938 bp, substantially larger than the GenomeScope haploid genome size estimate but in line with previous estimates of 2.04 Gbp [38]. The assembly had a scaffold N50 of 174.24 Mbp and a GC content of 35.62% [52] (Table 1).

Seven scaffolds in Vvill1.0 represented haploid representations of the seven estimated linkage groups for *V. villosa*, accounting for 67.74% of the total genome assembly size [52]. Roughly one-third of the assembly was impossible to assign to linkage group scaffolds due to the high heterozygosity of the sequenced individual. This high heterozygosity is probably due to the cross-pollinating nature of *V. villosa* in comparison to the more selfing nature of *V. sativa*. A total of 5373 contigs were obtained with an N50 of approximately 600 kbp (Table 1). The repetitive content in Vvill1.0 was similar to that of the *V. sativa* reference assembly, with 81.1% of the assembly consisting of identified repeats, compared to 83.9% in *V. sativa*. Vvill1.0 was predicted to have an LAI of 22.5, categorizing it as a high-quality reference genome in the "gold" category due to the assembly fidelity of repeat elements. Using prediction and RNA-seq evidence, the authors identified 53,321 protein-coding genes, with 81.8% of them annotated [52].

2.4. Lens culinaris and L. ervoides

Lentil (*Lens culinaris* Medik.) is a self-pollinated legume crop with a diploid chromosome number of 2n = 2x= 14. Originating from the Middle East and Central Asia, lentil is now cultivated worldwide and serves as an important source of plant-based protein [53]. It is a dietary staple in many Middle Eastern and South Asian countries and is gaining popularity in other regions due to its cooking flexibility and suitability for diabetic, gluten-free, and heart-smart diets [15].

To improve lentil's important traits like disease resistance, environmental stress tolerance, and agronomic and quality traits, efforts have been made to introduce novel alleles from the wild gene pool into elite cultivated backgrounds. *Lens ervoides* (Bring.) Grande was initially used as a source of disease resistance, and it also offers variability for seed quality and yield traits [54,55].

In a recent study, Ramsay et al. [15] assembled the genomes of two lentil species: the Canadian red *L. culinaris* cultivar CDC Redberry and the wild *L. ervoides* accession IG 72815. The *L. culinaris* reference genome (Lcu.2RBY; CDC Redberry) was assembled using 54x long-reads, polished with both Oxford Nanopore long reads and additional Illumina short reads. Hi-C proximity and a single genetic map were used to scaffold contigs into seven pseudomolecules, covering 92.8% of the 3.92 Gbp estimated genome size (Table 1).

The *L. ervoides* reference genome (Ler.1DRT; IG 72815) was assembled from 52x long reads, polished with long reads and additional short reads, and similarly scaffolded into seven pseudomolecules, representing 96.1% of the assembly. This assembly is smaller, approximately 2.9 Gbp [15] (Table 1).

Both genomes have a similar number of high-confidence genes, with 39,778 in Lcu.2RBY and 37,045 in Ler.1DRT, representing, respectively, 94% and 95% complete BUSCO with the Fabales database (Table 1). Repeats constitute a significant portion of the CDC Redberry genome, accounting for 82.6% (3.2 Gbp) of the 3.9 Gbp assembly. In IG 72815, the proportion of repeats is lower at 78.3%, with 947 Mbp fewer LTRs, contributing to the smaller genome size. Like the other species from the Fabeae tribe, the repetitive regions primarily consist mostly of LTRs of the *Ty3*/*gypsy* type. RNA-seq data from seven tissue samples: flowers, leaflets, seedling roots, root nodules, seedlings, etiolated seedlings, and flower buds was used to assist with the annotation of protein-coding genes in

L. culinaris. For *L. ervoides*, RNA-seq data from Ascochyta lentis-inoculated samples were used for annotation [15].

This study also found high collinearity among pea, Medicago, and the two lentil species, with some notable translocations and inversions. A large, inverted, reciprocal translocation exists between chromosomes 2 and 3 in both *Lens* species relative to pea chromosomes 1 and 5, suggesting this is a specific feature of the pea genome [15]. Chromosome-level structural variations, such as translocations and inversions, are considered drivers of genome evolution and speciation [56]. Comparative mapping indicated several rearrangements occurred after *Lens* species diverged from peas and other cool-season legumes but before *L. culinaris* and *L. ervoides* diverged [15].

Having genome assemblies for both wild and cultivated lines will facilitate more precise and cost-effective access to the genetic variability available in wild relatives. Mapping sequencing reads to their own genome instead of relying on alignment to a related domesticated species, resulting in more markers, improved QTL analyses, and a better chance of identifying candidate genes [57]. These findings will significantly aid in further research and genetic improvement of lentil as an important crop.

3. Applications in Abiotic and Biotic Studies

With the advances in sequencing and the availability of all these legume reference genomes, SNP discovery and QTL mapping can be performed, and QTLs can be more precisely localized on the genetic and physical sequence maps. This enables a better knowledge of the genetic basis of traits that are important to breeders, facilitating the generation of improved varieties. Nowadays, resistance or tolerance to biotic and abiotic stresses are among the most pressing traits to breed for. With accelerating climate change, increased abiotic stresses are expected to challenge agriculture and food security [58]. High-temperature spikes during crop growth, especially for the most critical reproductive period, are expected to exceed the range encountered during crop domestication [59]. Likewise, biotic stresses, such as pathogen infection and pest infestation, are also affected by changes in temperature and humidity. Warmer and more humid conditions tend to accelerate the development of pests, enabling their range expansion. Consequently, new pathogens and pests could emerge in areas where they have not previously occurred, affecting crops [60]. A summary of applications in biotic and abiotic stress studies using the genomic information from the Fabeae tribe is presented in Table 2.

Table 2. Summary of applications in biotic and abiotic stress studies using the genomic information from the Fabeae tribe. n.a. means not available.

Species	Biotic Stress			Abiotic Stress		
	Trait	**Approach**	**References**	**Trait**	**Approach**	**References**
P. sativum cv. Caméor	Powdery mildew resistance	GWAS and comparative genomics	[61–63]	Drought tolerance	Genotyping, GWAS, and proteomics	[64–68]
	White mold resistance	candidate genes identification	[69]	Heat tolerance	GWAS, candidate genes identification, and proteomics	[19,68,70]
	Rust resistance	GWAS and comparative genomics	[71,72]	Frost tolerance	Candidate genes identification and syntenic studies	[73,74]
	Ascochyta blight resistance	candidate genes identification		Cold tolerance	Mapping of transcriptome and miRNA sequencing	[75]

	Fusarium wilt and root rot resistance	GWAS, QTL mapping, and candidate gene identification	[76,77]	Salinity tolerance	Mining ortholog genes and miRNAs	[78]
	Aphanomyces root rot resistance	SNP quality assessment, QTL mapping, and proteomics	[79–81]			
	Aphid resistance	GWAS	[82]			
	Bruchid resistance	QTL mapping and candidate gene identification	[83]			
	Parasitic plant Orobanche crenata resistance	Comparative mapping	[20]			
	Herbivore resistance	Candidate gene identification	[84]			
V. faba cv. Hedin/2		n.a.		Drought tolerance	Transcriptome analysis	[85]
				Drought and salt stress	Gene expression	[86]
				Drought tolerance (PRO and SPAD levels)	GWAS	[87]
L. culinaris	Ascochyta blight resistance	Transcriptome (RNA-seq) analysis	[88]	Salt stress	GWAS	[89]
	Anthracnose race 1 resistance	GWAS	[90]			
	Stemphylium blight resistance	GWAS	[91]			
	Powdery mildew (MLO genes)	GWAS	[92]			

Pea crop suffers from various biotic stresses, including bacterial, fungal, viral, pest, and nematode infestations, as well as weeds. Among them, fungal diseases are considered to be the major type of biotic stress [2]. After the publication of the first pea genome [13], several studies on legume biotic stresses have been favored by the availability of a pea reference genome. Genetic and genomic studies on legume fungal diseases have been performed in aerial fungal diseases, such as powdery mildew [61–63], white mold [69], rust [71,72], ascochyta blight [93,94], and also in root diseases, like fusarium wilt and root rot [76,77,95] and Aphanomyces root rot (caused by oomycetes) [79–81]. In these studies, the Caméor pea genome was used for SNP calling, location of SNP physical positions, identification of candidate genes underlying QTLs, and syntenic studies aiming at genetic linkage maps validation and location of ortholog genes of interest by comparative mapping.

Comparative mapping between *L. sativus* and *L. cicera* genetic linkage maps and *P. sativum* cv. Caméor, *L. culinaris* Medik., and *M. truncatula* Gaertn. reference genomes revealed important aspects of the conservation of the *MLO1* powdery mildew susceptibility gene position and of the overall chromosomal rearrangements occurring during legume

evolution [61]. In this study, high synteny between grass pea recombinant inbred line high-density linkage map and the pea genomes was observed. Since the *L. sativus* genome was not fully assembled before 2023, comparative mapping between *Lathyrus* species and pea cv. Caméor genome [13] was also used for GWAS, in which SNP markers' physical positions were assigned based on the pea genome [61,63]. Furthermore, candidate genes underlying molecular markers identified by GWAS and QTL linkage mapping related to powdery mildew, rust, and Fusarium wilt resistance in *Lathyrus* spp. were predicted by comparative mapping using pea cv. Caméor genome [61–63,71,72,76]. Therefore, the availability of the pea genome allows the implementation of more accurate comparative genomics, reinforcing breeding programs through faster and more efficient identification of QTLs and candidate genes. Accordingly, SNP quality assessment and QTL location for partial resistance to *Aphanomyces root rot* were performed by alignments to the genome. Molecular markers identified as associated with pea resistance; for example, *Sclerotinia sclerotiorum* and *P. pinodes* were aligned to the pea genome for candidate gene discovery [69].

Proteomic studies also used pea cv. Caméor genome for identification of proteases secreted by the root pathogen *Aphanomyces euteiches* during pea infection, using comparison of MS/MS spectral against the reference genome [80,96].

Other studies on biotic stresses have been taking advantage of the Caméor pea genome availability, such as studies on resistance against insects [82,83], the parasitic plant *Orobanche crenata* [20], herbivores [84], and bacterial blight [97]. In these studies, the pea genome was useful for genotyping using the genome as a reference in GWAS for pea resistance to aphids [82]; comparative mapping between a high-density genetic map and the pea genome for genetic dissection of pea resistance to broomrape (*Orobanche crenata*) [20]; identification of candidate genes related to pea response to bruchid and herbivores [83,84]; and prediction of grain yield for field pea using bacterial blight disease scores [97]. The ZW6 pea genome was used in a comprehensive dataset of full- and partial-length NLR (nucleotide-binding leucine-rich repeat immune receptors) resistance genes across 100 chromosome-level plant genomes [98]. Plant NLR proteins directly or indirectly detect pathogen effectors activating the effector-triggered immunity (ETI), a robust layer of plant defense against pathogens [99].

Peas are also highly susceptible to abiotic stresses such as drought, heat, and cold conditions. Globally, drought and heat stresses are the most critical pea abiotic stresses, mainly during the flowering stage, while frost, salinity, and early-season flooding diversely affect the growth and yield of the crop. Fundamental knowledge and molecular tools for breeding have been boosted on the legume tolerance against abiotic stresses after peas entered the genomic era with the pea Caméor genome sequenced. Diverse studies have been performed to better understand the legume tolerance mainly to drought [64–68] and heat [19,68,70] stresses but also to frost [73,74], cold [75], and salinity [78]. For drought, the pea cv. Caméor genome has been used for genotyping and GWAS, for example, in the identification of invertases in seeds subjected to drought and protein identification in root nodules under drought conditions [64]. For heat stress, researchers have been using pea cv. Caméor genome for GWAS and further identification of candidate genes for heat adaptative traits [68] and mapping of transcriptomic reads related to the negative effects of heat stress on reproductive growth [70,75]. Similar approaches have been applied to study pea tolerance to frost, including the identification of candidate genes for this trait [73]. Syntenic studies were also performed between QTLs identified in faba bean for frost tolerance and the pea cv. Caméor genome and conservation of genetics for this trait were observed [74]. Mapping of transcriptome and miRNA sequencing reads against pea cv. Caméor genome revealed a microRNA regulation network involved in pea response to cold [75]. Finally, one study on pea photosynthetic performance under salinity conditions used pea cv. Caméor genome for mining ortholog genes and miRNAs of interest in *A. thaliana* [78].

The other pea assembly genomes available (at scaffold and contig levels) were not used for studies on biotic and abiotic stresses .

Regarding the *Lathyrus* genome assemblies, no reports on biotic or abiotic studies have been published so far. Nevertheless, the *L. sativus* Rbp assembly was used to study the macroevolution of NLR genes in legumes, including *L. sativus* [100]. Studies using *L. sativus* genome assemblies are focused on the biosynthetic pathways of B-ODAP. Nevertheless, due to the known enhanced resistance and tolerance of *L. sativus* against biotic and environmental conditions compared to other members of the Fabeae tribe, a reliable and accurate genome assembly will be crucial to understanding the *L. sativus* mechanisms underlying its response to stresses. Moreover, the development of new precision breeding screening tools will enable more competitive stress-resistant/tolerant *Lathyrus* and other legume varieties to be selected with increased speed and efficiency, accelerating their future introduction in the market.

The availability of the faba bean genome sequencing will greatly benefit breeders and geneticists, helping them improve sustainable protein production in the Mediterranean, subtropical, and northern temperate zones. It is known that faba bean production is severely affected by drought and cold [51,101,102], foliar diseases such as chocolate spot, Ascochyta blight, and Cercospora leaf spot, insects' pests such as broad-bean beetles (*Bruchus rufimanus*) and aphids (*Megoura crassicauda, Aphis fabae* Scop.), viral diseases including broad bean mottle virus (BBMV), broad bean stain virus (BBSV), bean leaf roll virus (BLRV), and bean yellow mosaic virus (BYMV) [103] and root-parasitic plants, such as broomrape (*Orobanche crenata*) [104]. Since the publication of *V. faba* genome assembly is still very recent, few reports on biotic and abiotic stresses have been published using the faba bean genome. The exceptions are three faba bean drought response studies [85,86]. The first study revealed that VfbZIP5 enhanced drought tolerance by reducing the levels of proline (PRO), malondialdehyde (MDA), and peroxidase (POD) [85]. In the second study, it was shown that the expression of *VfNR1* is a consequence of hypoxia in the root nodules that occurs in the presence of drought and salt stress [86]. The third is a genome-wide association study (GWAS), in which the faba bean genome, together with the *M. truncatula* genome, were used to annotate the potential candidate genes identified associated with seed weight, SPAD, and proline content under drought stress [87]. Simultaneously, traits related to seed quality and consumer preferences, such as low alkaloid glycosides vicine and convicine and low tannin content, are expected to be more efficiently targeted for faba bean breeding [105,106]. We foresee, in the near future, the publication of genome-wide association studies aiming at the identification of novel genetic variants and candidate genes controlling the above-mentioned important traits. As a practical application of the genome sequence, the genetic basis of seed size and hilum color was already dissected using high-resolution GWAS, highlighting the role of *Vfaba.Hedin2.R1.4g051440* gene in regulating seed size and polyphenol oxidase (PPO) in hilum color variation [11]. Additionally, the introduction of favorable alleles or the genomic edition of elite cultivars to speed up faba bean breeding and obtain new varieties with desirable qualities might be a reality soon. The release of the genome sequence also enables population-scale resequencing of genebank collections, elite cultivars, and mutants, further enhancing research and progress in this vital crop [106].

Moreover, the crucial insights into the genome characteristics of common vetch and hairy vetch offer valuable information to design whole genome sequencing strategies for crop improvement in the future. The availability of chromosome-scale genome assembly for both *V. sativa* and *V. villosa* is a significant development for genetic improvement programs targeting these crucial cover crops and forage species. The genomic information available can now assist the breeding for drought and cold tolerance [107,108], and lower content of γ-glutamyl-β-cyano-alanine (GBCA) toxin in common vetch [109], and for *Ramularia sphaeroidea* Sacc. and *Ascochyta viciae* resistance [110] and cold tolerance [111] in *V. villosa*. Additionally, since these species share conserved gene orthologs, they are valuable resources for comparative studies of leguminous plant genomes.

In lentil, the use of a draft version of the *L. culinaris* reference genome was already useful to improve the quality of an RNA-seq analysis of a defense transcriptome activated by *A. lentis* infection, in comparison to the de novo assembly of the transcriptome [88]. Also, recent genome-wide association studies benefited from the disclosure of the lentil genome, allowing the unveiling of potential candidate genes in salt tolerance mechanisms, the identification of anthracnose race 1 and stemphylium blight resistance loci, and lentil *MLO* genes [89–92].

To our knowledge, there is no report on the use of *L. ervoides* genome sequence in abiotic or biotic studies. However, this wild relative of lentil has shown tolerance against drought and cold stresses [112] and resistance to key biotic stresses, such as rust (*Uromyces fabae* (Grev.) Fuckel), fusarium wilt (*Fusarium oxysporum* f. sp. *lentis* (Vasd. Srin.) Gord), and powdery mildew (*Erysiphe trifolii*) [113]. While crossing between these two species is possible, *L. ervoides* is an accessible and promising source of useful genes/alleles for gene introgression aiming at widening the genetic base of lentils through interspecies crossing [15].

Transposons mediated unequal recombination, which potentially gives rise to the gain or loss of genes in legumes, especially in Fabeae, due to the high number of repetitive elements in this tribe [114]. Differential expansion and removal of these elements probably shaped genomes throughout the evolution of the Fabeae and notably within *Pisum*, highlighting the repetitive elements as a major driver in the evolution of these large genomes. Additionally, WGD also plays a significant role in plant genome evolution by providing raw materials that can be modified by natural or artificial selection.

Despite Fabeae members having a large size genome, some species reveal an overall contraction of the NLRome, which includes genes involved in plant defense mechanisms [100]. *P. sativum* and *L. sativus* have shown the lowest NLR density concerning their genome size. Contrary to this contraction trend, the legume tribe *Trifolieae* has shown large-scale expansion of NLRome irrespective of their genome size. These data suggest that the expansion and contraction of NLRs are not dependent on genome size [100].

Now that more genomic tools are available, we foresee an increase in the development of functional markers to assist crop breeding by marker-assisted selection. Several examples of functional markers, including cleaved amplified polymorphic sequence (CAPS), co-dominant, and KASPar markers, have been developed for the selection of er1 powdery mildew (*E. pisi*) resistance in pea cultivars [115–118]. In lentils, a genetic linkage map between *L. culinaris* and *L. odemensis* was developed based on thousands of functional markers (SNPs) mapped in expressed genes for morphological and agronomical traits and QTLs involved in resistance to *Ascochyta* [119]. Functional genomics approaches are crucial to introduce QTLs/genes associated with biotic or abiotic traits in desirable cultivars in a more efficient and accurate way. Syntenic regions shared between the genomes of different legume species will facilitate future comparative genomic studies.

4. Conclusions and Future Directions

The sequencing of legume genomes has grown very fast in the last decade, leading to an increase in the quantity and quality of publicly available genomic resources. The more advanced, accurate, and fast sequencing technologies, such as long-read sequencing and new computational tools, combined with the declining sequencing costs, have enabled the assembly of the large and complex genome characteristics of legumes. Complete genome assemblies will increase the knowledge of the evolution of the different species. This will have a tremendous impact on the understanding of the genetics and genomics of important agronomic traits, such as biotic and abiotic stress responses, and support legume breeding.

The proportions of repeats (78–85%) of the Fabeae genome assemblies described in this review are much higher than for other diploid legumes, such as Medicago (20–23%), chickpea (*Cicer arietinum* L.) (49%), pigeonpea (*Cajanus cajan* (L.) Huth.) (52%), and common bean (*Phaseolus vulgaris* L.) (45%). This helps to explain the larger size of these

genomes in comparison to other diploid plants, but also the rearrangement observed among some of these species, which promoted genome evolution and speciation.

A challenge for the generation of high-quality genome assemblies is the presence of high levels of heterozygosity, observed, for instance, in predominantly outcrossing species like *V. villosa* and *L. tuberosus*. High heterozygosity constitutes an important source of assembly error hindering the generation of high-quality reference genomes. This issue might be overcome by using an approach that includes different strategies, such as using longer reads, the combination of more than one assembly technique, both manual and automated data curation, and using gamete binning, a method that allows the separation of the whole-genome sequencing reads into haplotype-specific reads sets [120–122].

The growing affordability of long-read DNA sequencing will soon contribute to the creation of pangenomes, incorporating all the genetic diversity within the species belonging to this tribe. The availability of more reference genomes enhances the quality of comparative genomics studies by improving genome synteny analysis [61,123]. Moreover, there is even potential for a more expansive super-pangenome that encompasses not only the species belonging to the Fabeae tribe but also closely related crops [124]. By comprehending this range of genetic differences, breeders can effectively exploit specific genome regions from distantly related wild species, which confer traits like disease resistance and adaptability to different climates. This valuable resource would shed light on the evolutionary journey of legumes and the genomes of their members.

Author Contributions: All authors contributed to literature revision and writing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fundação para a Ciência e Tecnologia (FCT, Portugal) through the R&D Research Unit GREEN-IT—Bioresources for Sustainability (UIDB/04551/2020 DOI 10.54499/UIDB/04551/2020 and UIDP/04551/2020 DOI 10.54499/UIDP/04551/2020) and the LS4FUTURE Associated Laboratory (LA/P/0087/2020). CS and STL were also supported by FCT through the Research Contracts by Stimulus of Scientific Employment, Individual Support 2017 and 2022 programs, references 2017.00198.CEECIND (https://doi.org/10.54499/CEECIND/00198/2017/CP1428/CT0002)and 2022.00163.CEECIND (https://doi.org/10.54499/2022.00163.CEECIND/CP1725/CT0021) respectively.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Baloglu, M.C.; Celik Altunoglu, Y.; Baloglu, P.; Yildiz, A.B.; Türkölmez, N.; Özden Çiftçi, Y. Gene-editing technologies and applications in legumes: Progress, evolution, and future prospects. *Front. Genet.* **2022**, *13*, 859437. https://doi.org/10.3389/fgene.2022.859437.
2. Rubiales, D.; Fondevilla, S.; Chen, W.; Gentzbittel, L.; Higgins, T.J.V.; Castillejo, M.A.; Singh, K.B.; Rispail, N. Achievements and challenges in legume breeding for pest and disease resistance. *Crit. Rev. Plant Sci.* **2015**, *34*, 195–236. https://doi.org/10.1080/07352689.2014.898445.
3. Araújo, S.S.; Beebe, S.; Crespi, M.; Delbreil, B.; González, E.M.; Gruber, V.; Lejeune-Henaut, I.; Link, W.; Monteros, M.J.; Prats, E.; et al. Abiotic stress responses in Legumes: Strategies used to cope with environmental challenges. *Crit. Rev. Plant Sci.* **2015**, *34*, 237–280. https://doi.org/10.1080/07352689.2014.898450.
4. Sham, A.; Moustafa, K.; Al-Ameri, S.; Al-Azzawi, A.; Iratni, R.; AbuQamar, S. Identification of Arabidopsis candidate genes in response to biotic and abiotic stresses using comparative microarrays. *PLoS ONE* **2015**, *10*, e0125666. https://doi.org/10.1371/journal.pone.0125666.
5. Egan, A.N.; Vatanparast, M. Advances in legume research in the genomics era. *Aust. Syst. Bot.* **2019**, *32*, 459–483. https://doi.org/10.1071/SB19019.
6. Schmutz, J.; Cannon, S.B.; Schlueter, J.; Ma, J.; Mitros, T.; Nelson, W.; Hyten, D.L.; Song, Q.; Thelen, J.J.; Cheng, J.; et al. Genome sequence of the palaeopolyploid soybean. *Nature* **2010**, *463*, 178–183.
7. Young, N.D.; Debellé, F.; Oldroyd, G.E.D.; Geurts, R.; Cannon, S.B.; Udvardi, M.K.; Benedito, V.A.; Mayer, K.F.X.; Gouzy, J.; Schoof, H.; et al. The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature* **2011**, *480*, 520–524. https://doi.org/10.1038/nature10625.
8. Schmutz, J.; McClean, P.E.; Mamidi, S.; Wu, G.A.; Cannon, S.B.; Grimwood, J.; Jenkins, J.; Shu, S.; Song, Q.; Chavarro, C.; et al. A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* **2014**, *46*, 707–713. https://doi.org/10.1038/ng.3008.

9. Kang, Y.J.; Kim, S.K.; Kim, M.Y.; Lestari, P.; Kim, K.H.; Ha, B.-K.; Jun, T.H.; Hwang, W.J.; Lee, T.; Lee, J.; et al. Genome sequence of mungbean and insights into evolution within *Vigna* species. *Nat. Commun.* **2014**, *5*, 5443. https://doi.org/10.1038/ncomms6443.
10. Varshney, R.K.; Chen, W.; Li, Y.; Bharti, A.K.; Saxena, R.K.; Schlueter, J.A.; Donoghue, M.T.A.; Azam, S.; Fan, G.; Whaley, A.M.; et al. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* **2012**, *30*, 83–89. https://doi.org/10.1038/nbt.2022.
11. Jayakodi, M.; Golicz, A.A.; Kreplak, J.; Fechete, L.I.; Angra, D.; Bednář, P.; Bornhofen, E.; Zhang, H.; Boussageon, R.; Kaur, S.; et al. The giant diploid faba genome unlocks variation in a global protein crop. *Nature* **2023**, *615*, 652–659. https://doi.org/10.1038/s41586-023-05791-5.
12. Yang, T.; Liu, R.; Luo, Y.; Hu, S.; Wang, D.; Wang, C.; Pandey, M.K.; Ge, S.; Xu, Q.; Li, N.; et al. Improved pea reference genome and pan-genome highlight genomic features and evolutionary characteristics. *Nat. Genet.* **2022**, *54*, 1553–1563. https://doi.org/10.1038/s41588-022-01172-2.
13. Kreplak, J.; Madoui, M.-A.; Cápal, P.; Novák, P.; Labadie, K.; Aubert, G.; Bayer, P.E.; Gali, K.K.; Syme, R.A.; Main, D.; et al. A reference genome for pea provides insight into legume genome evolution. *Nat. Genet.* **2019**, *51*, 1411–1422. https://doi.org/10.1038/s41588-019-0480-1.
14. Shirasawa, K.; Sasaki, K.; Hirakawa, H.; Isobe, S. Genomic region associated with pod color variation in pea (*Pisum sativum*). *G3* **2021**, *11*, jkab081.
15. Ramsay, L.; Koh, C.S.; Kagale, S.; Gao, D.; Kaur, S.; Haile, T.; Gela, T.S.; Chen, L.-A.; Cao, Z.; Konkin, D.J.; et al. Genomic rearrangements have consequences for introgression breeding as revealed by genome assemblies of wild and cultivated lentil species. *bioRxiv* 2021, *preprint*. https://doi.org/10.1101/2021.07.23.453237.
16. Leitão, S.T.; Santos, C.; Araújo, S.d.S.; Rubiales, D.; Vaz Patto, M.C. Shared and tailored common bean transcriptomic responses to combined fusarium wilt and water deficit. *Hortic. Res.* **2021**, *8*, 149. https://doi.org/10.1038/s41438-021-00583-2.
17. Chen, Z.; Ly Vu, J.; Ly Vu, B.; Buitink, J.; Leprince, O.; Verdier, J. Genome-wide association sudies of seed performance traits in response to heat stress in Medicago truncatula uncover MIEL1 as a regulator of seed germination plasticity. *Front. Plant Sci.* **2021**, *12*, 673072. https://doi.org/10.3389/fpls.2021.673072.
18. Ji, J.; Zhang, C.; Sun, Z.; Wang, L.; Duanmu, D.; Fan, Q. Genome editing in cowpea *Vigna unguiculata* using CRISPR-Cas9. *Int. J. Mol. Sci.* **2019**, *20*, 2471.
19. Huang, S.; Gali, K.K.; Arganosa, G.C.; Tar'an, B.; Bueckert, R.A.; Warkentin, T.D. Breeding indicators for high-yielding field pea under normal and heat stress environments. *Can. J. Plant Sci.* **2023**, *103*, 259–269. https://doi.org/10.1139/cjps-2022-0158.
20. Delvento, C.; Arcieri, F.; Marcotrigiano, A.R.; Guerriero, M.; Fanelli, V.; Dellino, M.; Curci, P.L.; Bouwmeester, H.; Lotti, C.; Ricciardi, L.; et al. High-density linkage mapping and genetic dissection of resistance to broomrape (*Orobanche crenata* Forsk.) in pea (*Pisum sativum* L.). *Front. Plant Sci.* **2023**, *14*, 1216297. https://doi.org/10.3389/fpls.2023.1216297.
21. Garg, V.; Dudchenko, O.; Wang, J.; Khan, A.W.; Gupta, S.; Kaur, P.; Han, K.; Saxena, R.K.; Kale, S.M.; Pham, M.; et al. Chromosome-length genome assemblies of six legume species provide insights into genome organization, evolution, and agronomic traits for crop improvement. *J. Adv. Res.* **2022**, *42*, 315–329. https://doi.org/10.1016/j.jare.2021.10.009.
22. Pellicer, J.; Leitch, I.J. The plant DNA C-values database (release 7.1): An updated online repository of plant genome size data for comparative studies. *New Phytol.* **2020**, *226*, 301–305. https://doi.org/10.1111/nph.16261.
23. FAIRsharing.org. Plant DNA C-Values Database (Plant C-Values). 2001. Available online: https://fairsharing.org/FAIRsharing.7qexb2 (accessed on 2 August 2023). https://doi.org/10.25504/FAIRsharing.7qexb2.
24. Schaefer, H.; Hechenleitner, P.; Santos-Guerra, A.; de Sequeira, M.M.; Pennington, R.T.; Kenicer, G.; Carine, M.A. Systematics, biogeography, and character evolution of the legume tribe Fabeae with special focus on the middle-Atlantic island lineages. *BMC Evol. Biol.* **2012**, *12*, 250. https://doi.org/10.1186/1471-2148-12-250.
25. Rajarammohan, S.; Kaur, L.; Verma, A.; Singh, D.; Mantri, S.; Roy, J.K.; Sharma, T.R.; Pareek, A.; Kandoth, P.K. Genome sequencing and assembly of *Lathyrus sativus*—A nutrient-rich hardy legume crop. *Sci. Data* **2023**, *10*, 32. https://doi.org/10.1038/s41597-022-01903-4.
26. Edwards, A.; Njaci, I.; Sarkar, A.; Jiang, Z.; Kaithakottil, G.G.; Moore, C.; Cheema, J.; Stevenson, C.E.M.; Rejzek, M.; Novák, P.; et al. Genomics and biochemical analyses reveal a metabolon key to β-L-ODAP biosynthesis in *Lathyrus sativus*. *Nat. Commun.* **2023**, *14*, 876. https://doi.org/10.1038/s41467-023-36503-2.
27. Flood, P.J.; Nowrousian, M.; Huettel, B.; Woehle, C.; Becker, K.; Wollenweber, T.E.; Begerow, D.; Grefen, C. Draft genome of the aardaker (*Lathyrus tuberosus* L.), a tuberous legume. *BMC Genom. Data* **2022**, *23*, 70. https://doi.org/10.1186/s12863-022-01083-5.
28. Ma, L.; Wang, X.; Yan, M.; Liu, F.; Zhang, S.; Wang, X. Genome survey sequencing of common vetch (*Vicia sativa* L.) and genetic diversity analysis of Chinese germplasm with genomic SSR markers. *Mol. Biol. Rep.* **2022**, *49*, 313–320. https://doi.org/10.1007/s11033-021-06875-z.
29. Shirasawa, K.; Kosugi, S.; Sasaki, K.; Ghelfi, A.; Okazaki, K.; Toyoda, A.; Hirakawa, H.; Isobe, S. Genome features of common vetch (*Vicia sativa*) in natural habitats. *Plant Direct* **2021**, *5*, e352. https://doi.org/10.1002/pld3.352.
30. Xi, H.; Nguyen, V.; Ward, C.; Liu, Z.; Searle, I.R. Chromosome-level assembly of the common vetch (*Vicia sativa*) reference genome. *GigaByte* **2022**, *2022*, gigabyte38.
31. Dahl, W.J.; Foster, L.M.; Tyler, R.T. Review of the health benefits of peas (*Pisum sativum* L.). *Br. J. Nutr.* **2012**, *108*, S3–S10.
32. Smýkal, P.; Varshney, R.K.; Singh, V.K.; Coyne, C.J.; Domoney, C.; Kejnovský, E.; Warkentin, T. From Mendel's discovery on pea to today's plant genetics and breeding. *Theor. Appl. Genet.* **2016**, *129*, 2267–2280. https://doi.org/10.1007/s00122-016-2803-2.

33. Macas, J.; Ávila Robledillo, L.; Kreplak, J.; Novák, P.; Koblížková, A.; Vrbová, I.; Burstin, J.; Neumann, P. Assembly of the 81.6 Mb centromere of pea chromosome 6 elucidates the structure and evolution of metapolycentric chromosomes. *PLoS Genet.* **2023**, *19*, e1010633. https://doi.org/10.1371/journal.pgen.1010633.
34. Amarasinghe, S.L.; Su, S.; Dong, X.; Dong, X.; Zappia, L.; Ritchie, M.E.; Gouil, Q. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol* **2020**, 21, 30. https://doi.org/10.1186/s13059-020-1935-5
35. Neumann, P.; Koblíz˘ková, A.; Navrátilová, A.; Macas, J. Significant expansion of *Vicia pannonica* genome size mediated by amplification of a single type of giant retroelement. *Genetics* **2006**, *173*, 1047–1056. https://doi.org/10.1534/genetics.106.056259.
36. Neumann, P.; Pozárková, D.; Macas, J. Highly abundant pea LTR retrotransposon Ogre is constitutively transcribed and partially spliced. *Plant Mol. Biol* .**2003**, *53*, 399–410.
37. Macas, J.; Neumann, P. Ogre elements—A distinct group of plant Ty3/gypsy-like retrotransposons. *Gene* **2007**, *390*, 108–116.
38. Macas, J.; Novák, P.; Pellicer, J.; Čížková, J.; Koblížková, A.; Neumann, P.; Fuková, I.; Doležel, J.; Kelly, L.J.; Leitch, I.J. In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the legume tribe Fabeae. *PLoS ONE* **2015**, *10*, e0143424.
39. Bayer, P.E.; Golicz, A.A.; Scheben, A.; Batley, J.; Edwards, D. Plant pan-genomes are the new reference. *Nat. Plants* **2020**, *6*, 914–920.
40. Emmrich, P.M.F.; Sarkar, A.; Njaci, I.; Kaithakottil, G.G.; Ellis, N.; Moore, C.; Edwards, A.; Heavens, D.; Waite, D.; Cheema, J.; et al. A draft genome of grass pea (*Lathyrus sativus*), a resilient diploid legume. *bioRxiv* 2020, *preprint*. https://doi.org/10.1101/2020.04.24.058164.
41. Lambein, F.; Travella, S.; Kuo, Y.-H.; Van Montagu, M.; Heijde, M. Grass pea (*Lathyrus sativus* L.): Orphan crop, nutraceutical or just plain food? *Planta* **2019**, *250*, 821–838. https://doi.org/10.1007/s00425-018-03084-0.
42. Vaz Patto, M.C.; Skiba, B.; Pang, E.C.K.; Ochatt, S.J.; Lambein, F.; Rubiales, D. *Lathyrus* improvement for resistance against biotic and abiotic stresses: From classical breeding to marker assisted selection. *Euphytica* **2006**, *147*, 133–147. https://doi.org/10.1007/s10681-006-3607-2.
43. Ellis, N.; Vaz Patto, M.C.; Rubiales, D.; Macas, J.; Novák, P.; Kumar, S.; Hao, X.; Edwards, A.; Sarkar, A.; Emmrich, P. Grasspea. In *Underutilised Crop Genomes*, Chapman, M.A., Ed.; Springer International Publishing: Cham, Switzerland, 2022; pp. 217–228.
44. Gonçalves, L.; Rubiales, D.; Bronze, M.R.; Vaz Patto, M.C. Grass pea (*Lathyrus sativus* L.): A Sustainable and resilient answer to climate challenges. *Agronomy* **2022**, *12*, 1324.
45. Vaz Patto, M.C.; Rubiales, D. Lathyrus diversity: Available resources with relevance to crop improvement—*L. sativus* and *L. cicera* as case studies. *Ann. Bot.* **2014**, *113*, 895–908. https://doi.org/10.1093/aob/mcu024.
46. Smykal, P.; Erdos, L. European tuberous *Lathyrus* species. *Legumes Perspect.—J. Int. Legume Soc.* **2020**, *19*, 36–38.
47. Zimin, A.V.; Puiu, D.; Luo, M.C.; Zhu, T.; Koren, S.; Marçais, G.; Yorke, J.A.; Dvořák, J.; Salzberg, S.L. Hybrid assembly of the large and highly repetitive genome of *Aegilops tauschii*, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm. *Genome Res.* **2017**, *27*, 787–792.
48. Cheng, H.; Concepcion, G.T.; Feng, X.; Zhang, H.; Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **2021**, *18*, 170–175. https://doi.org/10.1038/s41592-020-01056-5.
49. Han, S.; Sebastin, R.; Wang, X.; Lee, K.J.; Cho, G.-T.; Hyun, D.Y.; Chung, J.-W. Identification of *Vicia* species native to South Korea using molecular and morphological characteristics. *Front. Plant Sci.* **2021**, *12*, 608559. https://doi.org/10.3389/fpls.2021.608559.
50. Cernay, C.; Pelzer, E.; Makowski, D. A global experimental dataset for assessing grain legume production. *Sci. Data* **2016**, *3*, 160084. https://doi.org/10.1038/sdata.2016.84.
51. Lyu, J.I.; Ramekar, R.; Kim, J.M.; Hung, N.N.; Seo, J.S.; Kim, J.B.; Choi, I.Y.; Park, K.C.; Kwon, S.J. Unraveling the complexity of faba bean (Vicia faba L.) transcriptome to reveal cold-stress-responsive genes using long-read isoform sequencing technology. *Sci. Rep.* **2021**, *11*, 21094.
52. Fuller, T.; Koch, L.M.; Kucek, L.K.; Ali, S.; Mangelson, H.; Hernandez, T.; Smith, T.P.L.; Bickhart, D.M.; Riday, H.; Sullivan, M.L. A reference assembly for the legume cover crop, hairy vetch (*Vicia villosa*). *GigaByte* **2023**, *2023*, gigabyte98. https://doi.org/10.46471/gigabyte.98.
53. Smýkal, P.; Coyne, C.J.; Ambrose, M.J.; Maxted, N.; Schaefer, H.; Blair, M.W.; Berger, J.; Greene, S.L.; Nelson, M.N.; Besharat, N.; et al. Legume crops phylogeny and genetic diversity for science and breeding. *Crit. Rev. Plant Sci.* **2015**, *34*, 43–104. https://doi.org/10.1080/07352689.2014.897904.
54. Vail, S.; Strelioff, J.V.; Tullu, A.; Vandenberg, A. Field evaluation of resistance to *Colletotrichum truncatum* in *Lens culinaris*, *Lens ervoides*, and *Lens ervoides*×*Lens culinaris* derivatives. *Field Crops Res.* **2012**, *126*, 145–151. https://doi.org/10.1016/j.fcr.2011.10.002.
55. Tahir, M.; Lindeboom, N.; Båga, M.; Vandenberg, A.; Chibbar, R. Composition and correlation between major seed constituents in selected lentil (*Lens culinaris*. Medik) genotypes. *Can. J. Plant Sci.* **2011**, *91*, 825–835. https://doi.org/10.4141/cjps2011-010.
56. Zhang, L.; Reifová, R.; Halenková, Z.; Gompert, Z. How important are structural variants for speciation? *Genes* **2021**, *12*, 1084.
57. Jaganathan, D.; Bohra, A.; Thudi, M.; Varshney, R.K. Fine mapping and gene cloning in the post-NGS era: Advances and prospects. *Theor. Appl. Genet.* **2020**, *133*, 1791–1810.
58. Muluneh, M.G. Impact of climate change on biodiversity and food security: A global perspective—A review article. *Agric. Food Secur.* **2021**, *10*, 36. https://doi.org/10.1186/s40066-021-00318-5.

59. Coyne, C.J.; Kumar, S.; von Wettberg, E.J.B.; Marques, E.; Berger, J.D.; Redden, R.J.; Ellis, T.H.N.; Brus, J.; Zablatzká, L.; Smýkal, P. Potential and limits of exploitation of crop wild relatives for pea, lentil, and chickpea improvement. *Legume Sci.* **2020**, *2*, e36. https://doi.org/10.1002/leg3.36.
60. McDonald, B.A.; Stukenbrock, E.H. Rapid emergence of pathogens in agro-ecosystems: Global threats to agricultural sustainability and food security. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2016**, *371*, 20160026.
61. Santos, C.; Polanco, C.; Rubiales, D.; Vaz Patto, M.C. The *MLO1* powdery mildew susceptibility gene in *Lathyrus* species: The power of high-density linkage maps in comparative mapping and synteny analysis. *Plant Genome* **2021**, *14*, 7.
62. Santos, C.; Martins, D.; Rubiales, D.; Vaz Patto, M.C. Partial resistance against *Erysiphe pisi* and *E. trifolii* under different genetic control in *Lathyrus cicera*: Outcomes from a linkage mapping approach. *Plant Dis.* **2020**, *104*, 2875–2884.
63. Martins, D.; Santos, C.; Sampaio, A.M.; Rubiales, D.; Patto, M.C.V. *Lathyrus sativus* resistance against the existing and emerging pathogens *Erysiphe pisi* and *E. trifolii*: A case of commonalities or total discrepancy? *Phytopathology®* **2023**, *113*, 866–872. https://doi.org/10.1094/phyto-06-22-0227-fi.
64. Shumilina, J.; Gorbach, D.; Popova, V.; Tsarev, A.; Kuznetsova, A.; Grashina, M.; Dorn, M.; Lukasheva, E.; Osmolovskaya, N.; Romanovskaya, E.; et al. Protein glycation and drought response of pea (*Pisum sativum* L.) root nodule proteome: A proteomics approach. *Biol. Commun.* **2021**, *66*, 210–224. https://doi.org/10.21638/spbu03.2021.303.
65. Henriet, C.; Balliau, T.; Aimé, D.; Le Signor, C.; Kreplak, J.; Zivy, M.; Gallardo, K.; Vernoud, V. Proteomics of developing pea seeds reveals a complex antioxidant network underlying the response to sulfur deficiency and water stress. *J. Exp. Bot.* **2021**, *72*, 2611–2626.
66. Marsh, J.I.; Nestor, B.J.; Petereit, J.; Tay Fernandez, C.G.; Bayer, P.E.; Batley, J.; Edwards, D. Legume-wide comparative analysis of pod shatter locus PDH1 reveals phaseoloid specificity, high cowpea expression, and stress responsive genomic context. *Plant J.* **2023**, *115*, 68–80. https://doi.org/10.1111/tpj.16209.
67. Morin, A.; Kadi, F.; Porcheron, B.; Vriet, C.; Maurousset, L.; Lemoine, R.; Pourtau, N.; Doidy, J. Genome-wide identification of invertases in *Fabaceae*, focusing on transcriptional regulation of Pisum sativum invertases in seed subjected to drought. *Physiol. Plant.* **2022**, *174*, e13673. https://doi.org/10.1111/ppl.13673.
68. Tafesse, E.G.; Gali, K.K.; Lachagari, V.B.R.; Bueckert, R.; Warkentin, T.D. Genome-wide association mapping for heat and drought adaptive traits in pea. *Genes* **2021**, *12*, 1897.
69. Ashtari Mahini, R.; Kumar, A.; Elias, E.M.; Fiedler, J.D.; Porter, L.D.; McPhee, K.E. Analysis and identification of QTL for resistance to *Sclerotinia sclerotiorum* in pea (*Pisum sativum* L.). *Front. Genet.* **2020**, *11*, 587968. https://doi.org/10.3389/fgene.2020.587968.
70. Kaur, H.; Ozga, J.A.; Reinecke, D.M. Balancing of hormonal biosynthesis and catabolism pathways, a strategy to ameliorate the negative effects of heat stress on reproductive growth. *Plant Cell Environ.* **2021**, *44*, 1486–1503. https://doi.org/10.1111/pce.13820.
71. Santos, C.; Martins, D.C.; González-Bernal, M.J.; Rubiales, D.; Vaz Patto, M.C. Integrating phenotypic and gene expression linkage mapping to dissect rust resistance in chickling pea. *Front. Plant Sci.* **2022**, *13*, 837613. https://doi.org/10.3389/fpls.2022.837613.
72. Martins, D.C.; Rubiales, D.; Vaz Patto, M.C. Association mapping of *Lathyrus sativus* disease response to *Uromyces pisi* reveals novel loci underlying partial resistance. *Front. Plant Sci.* **2022**, *13*, 842545.
73. Beji, S.; Fontaine, V.; Devaux, R.; Thomas, M.; Negro, S.S.; Bahrman, N.; Siol, M.; Aubert, G.; Burstin, J.; Hilbert, J.-L.; et al. Genome-wide association study identifies favorable SNP alleles and candidate genes for frost tolerance in pea. *BMC Genom.* **2020**, *21*, 536. https://doi.org/10.1186/s12864-020-06928-w.
74. Carrillo-Perdomo, E.; Magnin-Robert, J.B.; Raffiot, B.; Deulvot, C.; Floriot, M.; Lejeune-Hénaut, I.; Marget, P.; Burstin, J.; Tayeh, N.; Aubert, G. A QTL approach in faba bean highlights the conservation of genetic control of frost tolerance among legume species. *Front. Plant Sci.* **2022**, *13*, 970865.
75. Mazurier, M.; Drouaud, J.; Bahrman, N.; Rau, A.; Lejeune-Hénaut, I.; Delbreil, B.; Legrand, S. Integrated sRNA-seq and RNA-seq analyses reveal a microRNA regulation network involved in cold response in *Pisum sativum* L. *Genes* **2022**, *13*, 1119.
76. Sampaio, A.M.; Rubiales, D.; Vaz Patto, M.C. Grass pea and pea phylogenetic relatedness reflected at *Fusarium oxysporum* host range. *Crop Prot.* **2021**, *141*, 105495. https://doi.org/10.1016/j.cropro.2020.105495.
77. Williamson-Benavides, B.A.; Sharpe, R.M.; Nelson, G.; Bodah, E.T.; Porter, L.D.; Dhingra, A. Identification of *Fusarium solani* f. sp. *pisi* (Fsp) responsive genes in *Pisum sativum*. *Front. Genet.* **2020**, *11*, 950. https://doi.org/10.3389/fgene.2020.00950.
78. Analin, B.; Mohanan, A.; Bakka, K.; Challabathula, D. Cytochrome oxidase and alternative oxidase pathways of mitochondrial electron transport chain are important for the photosynthetic performance of pea plants under salinity stress conditions. *Plant Physiol. Biochem.* **2020**, *154*, 248–259. https://doi.org/10.1016/j.plaphy.2020.05.022.
79. Wu, L.; Fredua-Agyeman, R.; Hwang, S.-F.; Chang, K.-F.; Conner, R.L.; McLaren, D.L.; Strelkov, S.E. Mapping QTL associated with partial resistance to *Aphanomyces* root rot in pea (*Pisum sativum* L.) using a 13.2 K SNP array and SSR markers. *Theor. Appl. Genet.* **2021**, *134*, 2965–2990. https://doi.org/10.1007/s00122-021-03871-6.
80. Kiselev, A.; Camborde, L.; Carballo, L.O.; Kaschani, F.; Kaiser, M.; van der Hoorn, R.A.L.; Gaulin, E. The root pathogen *Aphanomyces euteiches* secretes modular proteases in pea apoplast during host infection. *Front. Plant Sci.* **2023**, *14*, 1140101. https://doi.org/10.3389/fpls.2023.1140101.
81. Bari, M.A.A.; Fonseka, D.; Stenger, J.; Zitnick-Anderson, K.; Atanda, S.A.; Morales, M.; Worral, H.; Piche, L.; Kim, J.; Johnson, J.; et al. A greenhouse-based high-throughput phenotyping platform for identification and genetic dissection of resistance to *Aphanomyces* root rot in field pea. *Plant Phenome J.* **2023**, *6*, e20063. https://doi.org/10.1002/ppj2.20063.

82. Ollivier, R.; Glory, I.; Cloteau, R.; Le Gallic, J.F.; Denis, G.; Morlière, S.; Miteul, H.; Rivière, J.P.; Lesné, A.; Klein, A.; et al. A major-effect genetic locus, ApRVII, controlling resistance against both adapted and non-adapted aphid biotypes in pea. *Theor. Appl. Genet.* **2022**, *135*, 1511–1528.
83. Yan, J.; Chen, J.; Lin, Y.; Yuan, X.; Somta, P.; Zhang, Y.; Zhang, Z.; Zhang, X.; Chen, X. Mapping of quantitative trait locus reveals PsXI gene encoding xylanase inhibitor as the candidate gene for bruchid (*Callosobruch*us spp.) resistance in pea (*Pisum sativum* L.). *Front. Plant Sci.* **2023**, *14*, 1057577.
84. Basu, S.; Clark, R.E.; Bera, S.; Casteel, C.L.; Crowder, D.W. Responses of pea plants to multiple antagonists are mediated by order of attack and phytohormone crosstalk. *Mol. Ecol.* **2021**, *30*, 4939–4948. https://doi.org/10.1111/mec.16103.
85. Huang, L.-T.; Liu, C.-Y.; Li, L.; Han, X.-S.; Chen, H.-W.; Jiao, C.-H.; Sha, A.-H. Genome-wide identification of bZIP transcription factors in faba bean based on transcriptome analysis and investigation of their function in drought response. *Plants* **2023**, *12*, 3041.
86. Chammakhi, C.; Boscari, A.; Pacoud, M.; Aubert, G.; Mhadhbi, H.; Brouquisse, R. Nitric oxide metabolic pathway in drought-stressed nodules of faba bean (*Vicia faba* L.). *Int. J. Mol. Sci.* **2022**, *23*, 13057.
87. Gutiérrez, N.; Pégard, M.; Balko, C.; Torres, A.M. Genome-wide association analysis for drought tolerance and associated traits in faba bean (*Vicia faba* L.). *Front. Plant Sci.* **2023**, *14*, 1091875. https://doi.org/10.3389/fpls.2023.1091875.
88. Sari, E.; Bhadauria, V.; Ramsay, L.; Borhan, M.H.; Lichtenzveig, J.; Bett, K.E.; Vandenberg, A.; Banniza, S. Defense responses of lentil (*Lens culinaris*) genotypes carrying non-allelic ascochyta blight resistance genes to *Ascochyta lentis* infection. *PLoS ONE* **2018**, *13*, e0204124. https://doi.org/10.1371/journal.pone.0204124.
89. Dissanayake, R.; Cogan, N.O.I.; Smith, K.F.; Kaur, S. Application of genomics to understand salt tolerance in lentil. *Genes* **2021**, *12*, 332.
90. Gela, T.; Ramsay, L.; Haile, T.A.; Vandenberg, A.; Bett, K. Identification of anthracnose race 1 resistance loci in lentil by integrating linkage mapping and genome-wide association study. *Plant Genome* **2021**, *14*, e20131. https://doi.org/10.1002/tpg2.20131.
91. Adabor, S.; Gela, T.S.; Banniza, S. Mapping of genomic regions linked to stemphylium blight (*Stemphylium botryosum* Wallr.) resistance in lentil using linkage mapping and marker-trait association analysis. *Res. Sq.* 2022, *preprint*. https://doi.org/10.21203/rs.3.rs-2010961/v1.
92. Polanco, C.; Sáenz de Miera, L.E.; Bett, K.; Pérez de la Vega, M. A genome-wide identification and comparative analysis of the lentil MLO genes. *PLoS ONE* **2018**, *13*, e0194945.
93. Martins, L.B.; Balint-Kurti, P.; Reberg-Horton, S.C. Genome-wide association study for morphological traits and resistance to *Peryonella pinodes* in the USDA pea single plant plus collection. *G3 Genes|Genomes|Genet.* **2022**, *12*, jkac168. https://doi.org/10.1093/g3journal/jkac168.
94. Boutet, G.; Lavaud, C.; Lesné, A.; Miteul, H.; Pilet-Nayel, M.-L.; Andrivon, D.; Lejeune-Hénaut, I.; Baranger, A. Five regions of the pea genome co-control partial resistance to *D. pinodes*, tolerance to frost, and some architectural or phenological traits. *Genes* **2023**, *14*, 1399.
95. Wu, L.; Fredua-Agyeman, R.; Strelkov, S.E.; Chang, K.-F.; Hwang, S.-F. Identification of quantitative trait loci associated with partial resistance to Fusarium root rot and wilt caused by *Fusarium graminearum* in field pea. *Front. Plant Sci.* **2022**, *12*, 784593. https://doi.org/10.3389/fpls.2021.784593.
96. van West, P.; Appiah, A.A.; Gow, N.A.R. Advances in research on oomycete root pathogens. *Physiol. Mol. Plant Pathol.* **2003**, *62*, 99–113. https://doi.org/10.1016/S0885-5765(03)00044-4.
97. Zhao, H.; Pandey, B.R.; Khansefid, M.; Khahrood, H.V.; Sudheesh, S.; Joshi, S.; Kant, S.; Kaur, S.; Rosewarne, G.M. Combining NDVI and bacterial blight score to predict grain yield in field pea. *Front. Plant Sci.* **2022**, *13*, 923381. https://doi.org/10.3389/fpls.2022.923381.
98. Li, X.; Ma, L.; Wang, Y.; Ye, C.; Guo, C.; Li, Y.; Mei, X.; Du, F.; Huang, H. PlantNLRatlas: A comprehensive dataset of full- and partial-length NLR resistance genes across 100 chromosome-level plant genomes. *Front. Plant Sci.* **2023**, *14*, 1178069. https://doi.org/10.3389/fpls.2023.1178069.
99. Jones, J.D.G.; Dangl, J.L. The plant immune system. *Nature* **2006**, *444*, 323–329. https://doi.org/10.1038/nature05286.
100. Qureshi, F.; Mehmood, A.; Khan, S.A.; Bilal, M.; Urooj, F.; Alyas, M.; Ijaz, J.; Zain, M.; Noreen, F.; Rani, S.; et al. Macroevolution of NLR genes in family *Fabaceae* provides evidence of clade specific expansion and contraction of NLRome in Vicioid clade. *Plant Stress* **2023**, *10*, 100254. https://doi.org/10.1016/j.stress.2023.100254.
101. Khan, H.R.; Paull, J.G.; Siddique, K.H.M.; Stoddard, F.L. Faba bean breeding for drought-affected environments: A physiological and agronomic perspective. *Field Crops Res.* **2010**, *115*, 279–286. https://doi.org/10.1016/j.fcr.2009.09.003.
102. Muktadir, M.A.; Adhikari, K.N.; Merchant, A.; Belachew, K.Y.; Vandenberg, A.; Stoddard, F.L.; Khazaei, H. Physiological and biochemical basis of faba bean breeding for drought adaptation—A review. *Agronomy* **2020**, *10*, 1345.
103. Rubiales, D.; Khazaei, H. Advances in disease and pest resistance in faba bean. *Theor. Appl. Genet.* **2022**, *135*, 3735–3756. https://doi.org/10.1007/s00122-021-04022-7.
104. Fernández-Aparicio, M.; Flores, F.; Rubiales, D. The effect of *Orobanche crenata* infection severity in faba bean, field pea, and grass pea productivity. *Front. Plant Sci.* **2016**, *7*, 1409. https://doi.org/10.3389/fpls.2016.01409.
105. Lippolis, A.; Roland, W.S.U.; Bocova, O.; Pouvreau, L.; Trindade, L.M. The challenge of breeding for reduced off-flavor in faba bean ingredients. *Front. Plant Sci.* **2023**, *14*, 1286803. https://doi.org/10.3389/fpls.2023.1286803.
106. Maalouf, F.; Hu, J.; O'Sullivan, D.M.; Zong, X.; Hamwieh, A.; Kumar, S.; Baum, M. Breeding and genomics status in faba bean (*Vicia faba*). *Plant Breed.* **2019**, *138*, 465–473. https://doi.org/10.1111/pbr.12644.

107. Min, X.; Wang, Q.; Wei, Z.; Liu, Z.; Liu, W. Full-length transcriptional analysis reveals the complex relationship of leaves and roots in responses to cold-drought combined stress in common vetch. *Front. Plant Sci.* **2022**, *13*, 976094. https://doi.org/10.3389/fpls.2022.976094.
108. De la Rosa, L.; Zambrana, E.; Ramirez-Parra, E. Molecular bases for drought tolerance in common vetch: Designing new molecular breeding tools. *BMC Plant Biol.* **2020**, *20*, 71. https://doi.org/10.1186/s12870-020-2267-z.
109. Nguyen, V.; Riley, S.; Nagel, S.; Fisk, I.; Searle, I.R. Common vetch: A drought tolerant, high protein neglected leguminous crop with potential as a sustainable food source. *Front. Plant Sci.* **2020**, *11*, 818. https://doi.org/10.3389/fpls.2020.00818.
110. Renzi, J.P.; Garayalde, A.F.; Brus, J.; Pohankova, T.; Smýkal, P.; Cantamutto, M.A. Environmental and agronomic determinants of hairy vetch (*Vicia villosa* Roth) seed yield in rainfed temperate agroecosystems. *Eur. J. Agron.* **2023**, *147*, 126822. https://doi.org/10.1016/j.eja.2023.126822.
111. Renzi, J.P.; Chantre, G.R.; Smýkal, P.; Presotto, A.D.; Zubiaga, L.; Garayalde, A.F.; Cantamutto, M.A. Diversity of naturalized hairy vetch (*Vicia villosa* Roth) populations in central Argentina as a source of potential adaptive traits for breeding. *Front. Plant Sci.* **2020**, *11*, 189. https://doi.org/10.3389/fpls.2020.00189.
112. Asghar, M.J.; Hameed, A.; Rizwan, M.; Shahid, M.; Atif, R.M. Chapter 17—Lentil wild genetic resource: A potential source of genetic improvement for biotic and abiotic stress tolerance. In *Wild germplasm for Genetic Improvement in Crop Plants*; Azhar, M.T., Wani, S.H., Eds.; Academic Press: Cambridge, MA, USA, 2021; pp. 321–341.
113. Singh, M.; Kumar, S.; Basandrai, A.K.; Basandrai, D.; Malhotra, N.; Saxena, D.R.; Gupta, D.; Sarker, A.; Singh, K. Evaluation and identification of wild lentil accessions for enhancing genetic gains of cultivated varieties. *PLoS ONE* **2020**, *15*, e0229554. https://doi.org/10.1371/journal.pone.0229554.
114. Soltis, P.S.; Marchant, D.B.; Van de Peer, Y.; Soltis, D.E. Polyploidy and genome evolution in plants. *Curr. Opin. Genet. Dev.* **2015**, *35*, 119–125.
115. Sun, S.; Deng, D.; Duan, C.; Zong, X.; Xu, D.; He, Y.; Zhu, Z. Two novel er1 alleles conferring powdery mildew (*Erysiphe pisi*) resistance identified in a worldwide collection of pea (*Pisum sativum* L.) germplasms. *Int. J. Mol. Sci.* **2019**, *20*, 5071.
116. Ma, Y.; Coyne, C.J.; Main, D.; Pavan, S.; Sun, S.; Zhu, Z.; Zong, X.; Leitão, J.; McGee, R.J. Development and validation of breeder-friendly KASPar markers for er1, a powdery mildew resistance gene in pea (*Pisum sativum* L.). *Mol. Breed.* **2017**, *37*, 151. https://doi.org/10.1007/s11032-017-0740-7.
117. Sudheesh, S.; Lombardi, M.; Leonforte, A.; Cogan, N.O.I.; Materne, M.; Forster, J.W.; Kaur, S. Consensus genetic map construction for field pea (*Pisum sativum* L.), trait dissection of biotic and abiotic stress tolerance and development of a diagnostic marker for the er1 powdery mildew resistance gene. *Plant Mol. Biol. Report.* **2015**, *33*, 1391–1403. https://doi.org/10.1007/s11105-014-0837-7.
118. Pavan, S.; Schiavulli, A.; Appiano, M.; Miacola, C.; Visser, R.G.F.; Bai, Y.; Lotti, C.; Ricciardi, L. Identification of a complete set of functional markers for the selection of er1 powdery mildew resistance in Pisum sativum L. *Mol. Breed.* **2012**, *31*, 247–253. https://doi.org/10.1007/s11032-012-9781-0.
119. Polanco, C.; Sáenz de Miera, L.E.; González, A.I.; García, P.; Fratini, R.; Vaquero, F.; Vences, F.J.; Pérez de la Vega, M. Construction of a high-density interspecific (*Lens culinaris* × *L. odemensis*) genetic map based on functional markers for mapping morphological and agronomical traits, and QTLs affecting resistance to *Ascochyta* in lentil. *PLoS ONE* **2019**, *14*, e0214409. https://doi.org/10.1371/journal.pone.0214409.
120. Campoy, J.A.; Sun, H.; Goel, M.; Jiao, W.-B.; Folz-Donahue, K.; Wang, N.; Rubio, M.; Liu, C.; Kukat, C.; Ruiz, D.; et al. Gamete binning: Chromosome-level and haplotype-resolved genome assembly enabled by high-throughput single-cell sequencing of gamete genomes. *Genome Biol.* **2020**, *21*, 306. https://doi.org/10.1186/s13059-020-02235-5.
121. Asalone, K.C.; Ryan, K.M.; Yamadi, M.; Cohen, A.L.; Farmer, W.G.; George, D.J.; Joppert, C.; Kim, K.; Mughal, M.F.; Said, R.; et al. Regional sequence expansion or collapse in heterozygous genome assemblies. *PLoS Comput. Biol.* **2020**, *16*, e1008104.
122. Kronenberg, Z.N.; Rhie, A.; Koren, S.; Concepcion, G.T.; Peluso, P.; Munson, K.M.; Porubsky, D.; Kuhn, K.; Mueller, K.A.; Low, W.Y.; et al. Extended haplotype-phasing of long-read de novo genome assemblies using Hi-C. *Nat. Commun.* **2021**, *12*, 1935. https://doi.org/10.1038/s41467-020-20536-y.
123. Liu, C.; Wang, Y.; Peng, J.; Fan, B.; Xu, D.; Wu, J.; Cao, Z.; Gao, Y.; Wang, X.; Li, S.; et al. High-quality genome assembly and pan-genome studies facilitate genetic discovery in mung bean and its improvement. *Plant Commun.* **2022**, *3*, 100352. https://doi.org/10.1016/j.xplc.2022.100352.
124. Khan, A.W.; Garg, V.; Roorkiwal, M.; Golicz, A.A.; Edwards, D.; Varshney, R.K. Super-pangenome by integrating the wild side of a species for accelerated crop improvement. *Trends Plant Sci.* **2020**, *25*, 148–158.