Review

Frontiers in Dissecting and Managing *Brassica* Diseases: From Reference-Based RGA Candidate Identification to Building Pan-RGAomes

Yueqi Zhang [†], **William Thomas** [†] [ID], **Philipp E. Bayer** [ID], **David Edwards** [ID] **and Jacqueline Batley** *[ID]

School of Biological Sciences, The University of Western Australia, Perth 6009, Australia; 22078946@student.uwa.edu.au (Y.Z.); william.thomas@research.uwa.edu.au (W.T.); philipp.bayer@uwa.edu.au (P.E.B.); dave.edwards@uwa.edu.au (D.E.)

* Correspondence: jacqueline.batley@uwa.edu.au
† These authors contributed equally to this work.

check for
updates

Abstract: The *Brassica* genus contains abundant economically important vegetable and oilseed crops, which are under threat of diseases caused by fungal, bacterial and viral pathogens. Resistance gene analogues (RGAs) are associated with quantitative and qualitative disease resistance and the identification of candidate RGAs associated with disease resistance is crucial for understanding the mechanism and management of diseases through breeding. The availability of *Brassica* genome assemblies has greatly facilitated reference-based quantitative trait loci (QTL) mapping for disease resistance. In addition, pangenomes, which characterise both core and variable genes, have been constructed for *B. rapa*, *B. oleracea* and *B. napus*. Genome-wide characterisation of RGAs using conserved domains and motifs in reference genomes and pangenomes reveals their clustered arrangements and presence of structural variations. Here, we comprehensively review RGA identification in important *Brassica* genome and pangenome assemblies. Comparison of the RGAs in QTL between resistant and susceptible individuals allows for efficient identification of candidate disease resistance genes. However, the reference-based QTL mapping and RGA candidate identification approach is restricted by the under-represented RGA diversity characterised in the limited number of *Brassica* assemblies. The species-wide repertoire of RGAs make up the pan-resistance gene analogue genome (pan-RGAome). Building a pan-RGAome, through either whole genome resequencing or resistance gene enrichment sequencing, would effectively capture RGA diversity, greatly expanding breeding resources that can be utilised for crop improvement.

Keywords: *Brassica*; RGA; reference-based QTL mapping; GWAS; RGASeq; pan-RGAome

1. Agricultural Importance of *Brassica* Species

The *Brassica* genus belongs to the Brassicaceae family and *Brassica* crops are cultivated as oilseeds, vegetables, condiments and forages [1]. Commonly cultivated *Brassica* species include the diploids *B. rapa* (AA genome: 2n = 2× = 20), *B. nigra* (BB: 2n = 2× = 16) and *B. oleracea* (CC: 2n = 2× = 18), and the allotetraploid species *B. juncea* (AABB: 2n = 4× = 36), *B. napus* (AACC: 2n = 4× = 38) and *B. carinata* (BBCC: 2n = 4× = 34). The triangle of U summarises the genomic relationship of these three important amphidiploid hybrids and their diploid progenitor species [2]. *B. rapa* originated in the Mediterranean and central Asia and is commonly known as turnip and Chinese cabbage, but has subspecies also cultivated for fodder and rapeseed [3]. Black mustard (*B. nigra*) seed is commonly used as a spice and condiment [4], whilst *B. oleracea* is mainly used as a vegetable crop, including

brussels sprouts, cabbage, cauliflower, broccoli and Chinese kale [5]. *B. napus*, or canola, is the third most important oilseed crop globally following soybean and palm [6]. *B. juncea* is cultivated as an oilseed, spice, vegetable and fodder crop species [7] and *B. carinata* originated from the Ethiopian plateau and is an important condiment, vegetable and oilseed crop in northeast Africa [8].

Several diseases hinder the production of *Brassica* crops, for example, blackleg disease, caused by the dothideomycete fungus *Leptosphaeria maculans,* is a major constraint to canola production in Australia, Europe and America [6]. The disease causes 10–20% average annual yield losses in Australia, Canada and the UK [9–11]. In addition, serious epidemics have destroyed entire crops in regions of France in 2000 [12] and in the Lower Eyre Peninsula, Australia, in 2003 [13]. Stem rot caused by *Sclerotinia sclerotiorum* leads to yield losses in canola production worldwide [14]. An average sclerotinia stem rot incidence of 10–20% has been reported in Canadian provinces each year, with an incidence of 94% within an individual canola crop recorded in Saskatchewan in 2010 [14]. The disease can also cause up to 80% yield loss in China [15]. Clubroot, caused by *Plasmodiophora brassicae*, offers challenges to crop production worldwide [16]. The disease led to 50% yield loss in cruciferous crops in Canada [17] and 20–30% yield loss in China [18]. In addition, turnip mosaic virus, downy mildew caused by the oomycete pathogen *Hyaloperonospora brassicae*, fusarium wilt caused by the fungus *Fusarium oxysporum* f. sp. *conglutinans* and black rot caused by the bacteria *Xanthomonas campestris* pv. *campestris* are important pathogens which affect *Brassica* crops worldwide [1].

2. Quantitative and Qualitative Disease Resistance Mechanisms in Plants

Quantitative disease resistance is underpinned by multiple genes of small effect which provide incomplete or partial resistance in plants [19]. The pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) contributes to quantitative resistance [20]; transmembrane pattern recognition receptors (PRRs), including receptor-like kinase (RLKs) and receptor-like proteins (RLPs), are involved in PTI and recognise pathogen- or microbe-associated molecular patterns (PAMPs/MAMPs). PTI is also important in nonhost resistance whereby most plants are resistant to most microbial pathogens.

PRRs function collaboratively in mediating defence signalling [21]. Many PRRs rely on the regulatory protein brassinosteroid insensitive 1-associated receptor kinase 1 (BAK1) and others on the somatic embryogenesis receptor-like kinases (SERKs) [22]. Some RLKs, such as the RLK suppressor of BIR1 (SOBIR1), associate with RLPs to transmit the signal to downstream components [23]. Several receptor-like cytoplasmic kinases (RLCKs) link the PRR complex to downstream defences, triggering a number of cellular events including an increase of cytoplasmic calcium cation and anion effluxes, and extracellular alkalisation, the production of reactive oxygen species (ROS), and the activation of MAP kinase (MPK) cascades [24]. Some RLKs play a role in both defence and developmental pathways such as BAK1 [21]. Many pathogen effectors impede PRR functions by targeting microbial patterns, PRRs or components in the signalling cascade or interfere with RLKs that regulate plant growth and development to increase plant susceptibility.

Qualitative disease resistance is underpinned by major genes which provide complete or near-complete resistance in plants [19]. The effector-triggered immunity (ETI) response commonly forms the basis of qualitative resistance in plants [20]. ETI is responsible for host-mediated defence against adapted pathogens by recognising specific effector (avirulence) proteins secreted into host cytoplasm by the pathogen. Qualitative resistance genes typically encode nucleotide-binding-site leucine-rich repeat (NLR) proteins; however, RLK types have also been identified [19]. The recognition of effectors by NLRs also leads to the accumulation of ROS, activation of MAP kinase cascades, and defence gene expression [25]. In addition, NLR activation manifests as a hypersensitive response whereby localised cell death occurs at the site of penetration by biotrophic pathogens.

The most abundant resistance gene family encodes the intracellular receptor proteins NLR, which are functionally analogous to NOD-like receptors in the animal innate immune system [20]. Canonical NLR genes are comprised of a conserved central NB-ARC (nucleotide-binding adaptor

shared by APAF-1, certain R gene products and CED-4) domain, a leucine-rich repeat (LRR) domain, varying in the number of repeats, at the C-terminus and a TIR (homologous to the Drosophila TOLL and human interleukin-1 receptor) or CC (coiled-coil) or CCr (a zinc finger or resistance to powdery mildew 8 (RPW8)) motif at the N-terminus. Based on this N-terminus motif NLRs are categorised into three subclasses TIR-NBS-LRR (TNL) CC-NBS-LRR (CNL) and RPW8-NBS-LRR (RNL).

PRRs function not only in immune response, but also regulation of plant growth and development, reproduction and symbiosis [21]. Plant RLKs, analogous to animal receptor tyrosine kinases, contain an ectodomain (ECD), transmembrane domain and a cytoplasmic kinase domain, while RLPs lack the kinase domain [21]. The ECD domain is highly variable allowing the recognition of a wide range of ligands including steroids, peptides, polysaccharides and lipopolysaccharides. RLKs can be classified into RD (arginine I and aspartate (D) amino acid residues at the activation site of the kinase domain) and non-RD types based on the conservation of the arginine residue, among which non-RD kinases are mainly involved in plant immunity [26].

3. Genome/Pangenome-Wide RGA Prediction in *Brassicas*

The advance in next generation sequencing technologies have facilitated the construction of multiple genome assemblies for *B. rapa* [27,28], *B. oleracea* [28–31], *B. napus* [32–35] and *B. juncea* [36,37]. The concept of the pangenome was proposed as the new reference genome beyond a single reference for a species [38]. Pangenomes represent the species-wide genome diversity, which includes core genes present in all individuals and variable genes absent in some individuals. Pangenomes for *B. oleracea* [39] and *B. napus* [40] have also been constructed.

The availability of genome and pangenome assemblies of several *Brassica* species have facilitated the genetic study of disease resistance mechanisms. RGAs can be predicted bioinformatically based on their conserved domains and motifs [41]. A number of homology-based computational methods have been applied to annotate RGAs using conserved motifs, domains and recognised structural features [20]. These methods are based on BLAST similarity searches (Pfam and InterProScan), sequence alignment with resistance gene databases (PRGdb), searches for domain sequence homology with Hidden Markov Models (HMMER) or motif discovery (MEME). An integrative RGA prediction pipeline, RGAugury, was developed to firstly identify conserved domains and motifs including NB-ARC, LRR, transmembraI (TM), serine/threonine and tyrosine kinase (STTK), lysin motif (LysM), CC and TIR, and then classify these domains into four major RGA families; NBS-encoding, TM-CC, and membrane-associated RLP and RLK [41]. Genome-wide RGA repertoires characterised in *Brassica* genome and pangenome assemblies are comprehensively summarised in Table 1.

Table 1. Summary of the Genome-wide Prediction of nucleotide-binding-site leucine-rich repeat (NLR), receptor-like kinase (RLKs) and receptor-like proteins (RLPs) in *Brassica* Genome Assemblies.

Species.	Assemblies	NLR	RLK	RLP	Prediction Methods	References for RGA Prediction
B. rapa	BAC clones of *B. rapa* ssp. *pekinensis cv. Chiifu*	92	NA	NA	FGENESH program, BLAST and EST sequences	[42]
	Chiifu-401-42 v1	206	NA	NA	HMMER	[43]
	Chiifu-401-42 v1	204	NA	NA	BLAST and HMMsearch	[44]
	Chiifu-401-42 v1	249	NA	NA	MEME/MAST and InterProScan	[45]
	Chiifu-401-42 v1	202	NA	NA	HMMER for prediction and verified on the Pfam website.	[46]
	Chiifu-401-42 v1	NA	303	NA	Pfam database	[47]
B. oleracea	TO1000 v1	157	NA	NA	HMMER	[43]
	TO1000 v1	443	NA	NA	MEME/MAST and InterProScan	[45]
	TO1000 v1 and *capitata* homozygous line 02–12	146	NA	NA	HMMER for prediction and verified on the Pfam website.	[46]
	capitata homozygous line 02–12	138	NA	NA	HMMER	[48]
	TO1000 v2.1	439	822	159	The RGAugury pipeline	[5]
	Pangenome	556	901	213	The RGAugury pipeline	[5]
B. napus	Darmor-*bzh* v4.1	425	NA	NA	NA	[32]
	Darmor-*bzh* v4.1	641	NA	NA	MEME/MAST and InterProScan	[45]
	Darmor-*bzh* v4.1	464	NA	NA	HMMER for prediction and verified on the Pfam website.	[46]
	Darmor-*bzh* v4.1	NA	NA	184	MEME/MAST for prediction and verified with Pfam, TMHMM, Phobius and SMART searches	[49]
	Darmor-*bzh* v4.1	621	1497	273	The RGAugury pipeline	[50]
	Darmor-*bzh* v8.1	286	989	77		
	Tapidor v6.3	208	680	223		
	ZS11	566	1517	260		
	Pangenome based on Darmor-*bzh* v8.1	503	1098	148	The RGAugury pipeline	[51]
B. juncea	v1.5	289	1225	228	InterProScan and the RGAugury pipeline	[52,53]
	v1.5 Yang_et_al_2016	315	1085	191	The RGAugury pipeline	[50]

3.1. B. rapa

The first characterisation of NLRs in *B. rapa* was based on bacterial artificial chromosome (BAC) clones of *B. rapa* ssp. *pekinensis cv. Chiifu* covering around 19–27% of the genome and 47–66% of the euchromatic gene space. As an incomplete assembly was used, only 92 NLRs were identified, using a BLASTP search of the NBS domains queried with 207 *A. thaliana* NLR genes from the NIBLRRS server, and verified with sequences from 26,416 unigenes and 127,144 expressed sequence tag (EST) sequences of *B. rapa* ssp. *pekinensis* cv. Chiifu downloaded from the KBGP database (The Korea *B. rapa* Genome Project) [42]. In comparison, many more NLRs were predicted in the *Chiifu-401-42* v.1 genome assembly [43–46]. Three studies identified similar numbers of NLRs, 206 [43], 204 [44] and 202 [46], in the *B. rapa Chiifu-401-42* v.1 assembly using BLAST and/or HMMER search with the Pfam NB-ARC domain (PF00931). In contrast, more NLR genes (249) were predicted using MAST/MEME (Motif alignment and search tool/Multiple em for motif elicitation) and InterProScan analyses [45]. A total of 303 RLK genes were predicted in the *B. rapa Chiifu-401-42* v.1 genome using Pfam IDs retrieved by searching in the Pfam database with previously published *A. thaliana* LRR-RLK protein sequences [47].

3.2. B. oleracea

A total of 157 NLRs were identified in the *B. oleracea* TO1000 v1 genome assembly through searching with the Pfam NB-ARC domain (PF00931) using the HMMER V3.0 programme [43]. Using HMMER v3.1b2 based on a different cultivar; *B. oleracea capitata* homozygous line 02–12, 138 NLR genes were identified [48]. A similar number of NLRs (146) were identified in the TO1000 v1 and *capitata* homozygous line 02–12 assemblies using HMMER V3.0 [46]. Applying the MEME/MAST and InterProScan approach, many more NLR genes (443) were identified in the same genome assembly [45]. Based on the improved assembly TO1000 v2.1 and using the RGAugury pipeline version 2017-10-21, Bayer et al. [5] identified 439 NLR, 822 RLK and 159 RLP genes. RGAs had clustered arrangements. For example, 27 gene clusters harbouring 81 NLRs which represented 55.48% of the total NLRs were clustered in the *B. oleracea* TO1000 v1 and *capitata* homozygous line 02–12 assemblies [46].

In comparison with predictions based on single genome assemblies, more RGAs with diverse patterns are predicted in the *B. oleracea* pangenome which contains nine cultivars (broccoli, brussels sprouts, cabbage, cauliflower, kale and kohlrabi) and the wild relative *B. macrocarpa* [5]. There are variations in abundance of RGA and genetic diversities among cultivars. Among the total of 556 NLRs, 901 RLKs and 213 RLPs in the pangenome, TO1000 had the lowest number while the wild relative *B. macrocarpa* had the highest number of the three RGA classes. Of all the RGA genes, 1231 genes were core and 167 were variable.

3.3. B. napus

Among the *B. napus* assemblies which contain predicted NLRs, RLPs and RLKs, Darmor-*bzh* v4.1 has the highest number of RGAs; a total of 621 NLRs, 1497 RLKs and 273 RLPs were identified in the Darmor-*bzh* v4.1 assembly using the RGAugury pipeline [50]. Several other studies only predicted NLRs [32,45,46] or RLPs [49]. The variations of the number of RGAs predicted in the four studies compared with that of Tirnaz et al. [50] based on the same genome assembly was likely due to differences in the prediction methods. For example, Alamery et al. [45] deployed MEME/MAST and InterProScan approaches, Fu et al. [46] made the prediction using HMMER followed with verification on the Pfam website, and Stotz et al. [49] used the MEME/MAST approach for prediction and the putative RGAs were verified with Pfam, TMHMM, Phobius and SMART searches. Also applying the RGAugury pipeline, 566 NLRs, 1517 RLKs and 260 RLPs were predicted in the cultivar ZS11 assembly [50].

Fewer NLRs, RLPs and RLKs were predicted by Tirnaz et al. [50] (286 NLRs, 989 RLKs and 77 RLPs) and by Dolatabadian et al. [51] (293 NLRs, 1010 RLKs and 78 RLPs) in the Darmor-*bzh* v8.1 assembly compared to the Darmor-*bzh* v4.1 assembly (621 NLRs, 1497 RLKs and 273 RLPs) [50]. Also, far fewer NLRs, RLPs and RLKs were predicted in Tapidor v6.3 (208, 680 and 223) compared

to Darmor-*bzh* v4.1 [50]. This disparity is likely caused by differences in sequencing, assembly and repeat masking approaches [33,54]. Tapidor v6.3 was assembled using Illumina and mate pair reads and scaffolded by comparison to the Darmor-*bzh* v8.1 assembly and its genetic map. In comparison, the Darmor-*bzh* v4.1 assembly was constructed mainly from 454 (>700 bp) and Sanger reads [32]. Resistance genes that are fused with transposable elements could be masked, which may partially explain the difference in the number of NLRs identified [54].

In comparison with predictions based on single genome assemblies, more RGAs with diverse patterns are predicted in the *B. napus* pangenome which was constructed from 24 oilseeds, three fodders, two swedes, two vegetables and 19 synthetic accessions [51]. A total of 503 NLRs, 1098 RLKs and 148 RLPs were identified. Of all these genes, 996 were core and 753 were variable. In addition, the nonsynthetic lines had fewer RGAs (91.87 RGAs per line) than synthetic lines (156.94 RGAs per line).

3.4. B. juncea

A total of 289 NLRs were identified based on the presence of a NB-ARC domain (PF00931) with InterProScan v5.20–59.0 using Pfam 30.0 and coil in the *B. juncea* v1.5 assembly [53]. A total of 1225 RLKs and 228 RLPs were predicted using the RGAugury pipeline [52]. Tirnaz et al. [50] identified 315 NLRs, 1085 RLKs and 191 RLPs using the RGAugury pipeline in the *B. juncea* v1.5 assembly.

4. Reference-Based Candidate RGA Identification

The availability of *Brassica* reference genome assemblies has greatly facilitated QTL mapping and candidate disease resistance gene identification [55]. For example, in a QTL study for sclerotinia stem rot resistance in *B. napus*, the *Brassica* 60K Illumina Infinium array was used to genotype a double haploid (DH) F1 population of the cross between the resistant line J964 and susceptible line J902 [56]. The single nucleotide polymorphism (SNP) markers were anchored onto the *B. napus* Darmor-*bzh* v4.1 assembly and 17 QTL were present on six chromosomes (A02, A09, C02, C03, C04 and C06). The expression of the TNL, *BnaC03g05380D*, which was present in the SRC3a QTL on chromosome C3 was upregulated in resistant plants compared with susceptible plants.

In a QTL study for clubroot resistance in *B. rapa*, the F2 population of a cross between the susceptible Chinese cabbage and resistant turnip lines was genotyped using double digest restriction site-associated DNA sequencing (ddRAD-seq) and resulting SNP markers were mapped onto the *B. rapa Chiifu-401-42* v.1 assembly [57]. In this study, three NLRs were identified in the new locus, *CRs*, on chromosome A08. In another QTL study for *V. longisporum* resistance in *B. napus*, the *Brassica* 60K Illumina Infinium array was used to genotype an F1 population of the cross Express617 × R53 and a nested association mapping (NAM) panel with five subpopulations [58]. The obtained SNP and single nucleotide absence polymorphism (SNaP) markers were anchored onto the *B. napus* Darmor-*bzh* v4.1 assembly and QTL mapping and genome-wide association studies (GWAS) analysis were performed. QTL for *V. longisporum* resistance were identified on A03, A05, C05, C08 and C09 and seven NLRs within the QTL on C09 were identified.

Bayer et al. [5] located genetic markers for QTL which were linked with sclerotinia stem rot, black rot, and clubroot in the *B. oleracea* pangenome and the QTL harboured 28 RLKs. Similarly, genetic markers for the QTL for several qualitative blackleg resistance genes including *Rlm1*, *Rlm3*, *Rlm4*, *Rlm7*, *Rlm9*, *LepR1* and *LepR2* were physically located in the *B. napus* pangenome [51]. A total of 60 RGAs were identified in the combined QTL for *Rlm1*, *Rlm3*, *Rlm4*, *Rlm7* and *Rlm9* on chromosome A07, 14 RGAs were located within the *LepR1* QTL on chromosome A02, and 32 RGAs were located within the *LepR2* QTL on chromosome A10.

Sequence comparison of RGAs located within QTL which are associated with disease resistance contributes to candidate gene identification. The TNL allele in the *Crr1a* locus in the resistant line G004 conferred resistance to clubroot isolate Ano-01in *B. rapa* [59]. In comparison, the susceptible allele in the A9709 line encoded a truncated NL protein which lacked more than half of the TIR domain

and included 50 substitutions, two deletions, and two insertions. As another example, five NLRs were present in the QTL for TuRB07 which conferred resistance to an isolate of Turnip mosaic virus strain C4 on chromosome A06 in the *B. rapa Chiifu-401-42* v1 assembly [60]. Three NLRs, *Bra018810*, *Bra018834* and *Bra018835*, contained deletions and *Bra018862* was truncated in the resistant line VC1 and Chiifu reference, thus they were not functional. *Bra018863*, a strong candidate for TuRB07, encoded a full-length CNL and there were polymorphisms in the CC, NBS and C-terminal domains in comparison with the allele in the susceptible cultivar SR5.

RGA comparison also contributed to the identification of qualitative blackleg resistance genes. The blackleg resistance gene *Rlm2* was found to be an allelic variant of the *LepR3* RLP locus in *B. napus* [61]. The two *LepR3* lines, Hyola 60 and Surpass 400, contained an identical allele encoding a single-exon coding sequence (CDS) of 2556 bp. The three *Rlm2* lines, Glacier DH24287, Tapidor and Bristol carried identical alleles encoding a single-exon CDS of 2778 bp. Seven *lepR3*/*rlm2 B. napus* lines, Westar N-o-1, Polo, Columbus, Quantum, Ning You, *B. napus* var. *napobrassica* landrace Hvammsrofa and Topas DH16516 shared an allele encoding a single-exon CDS of 2853 bp. The *rlm2*/*lepR3* alleles in the *B. napus* line Marnoo and the *B. juncea* line AC Vulcan were two-exon CDS of 2219 bp and 2218 bp respectively. Two distinct susceptible alleles encoding single-exon CDS of 2916 bp and 2772 bp were identified in *B. rapa* subsp. *sylvestris* 006-1-1 and the *B. rapa* variety Torch respectively.

5. RGASeq for More Efficient Candidate Identification

The conventional approaches to identify causative RGAs are limited in their ability to precisely pinpoint the causative gene. QTL mapping often only enables the identification of a large genomic region which can contain multiple RGAs and is restricted to the genetic recombination present in a biparental mapping population. GWAS overcomes these limitations, but its reliance on a reference genome, particularly if only a single reference is used, means that significant RGA variation from the reference won't be detected [62]. These constraints result in extensive fine mapping and gene validation in order to confidently determine the causative RGA [63,64]. Resistance gene enrichment sequencing or RenSeq, is a targeted enrichment sequencing workflow, which enables the capture of plant NLR sequences via the hybridisation of RNA baits [65]. Targeted enrichment of RLPs and RLKs is through RLP/KSeq, a recent adaptation of RenSeq [66]. RenSeq and RLP/KSeq (hereafter collectively referred to as RGASeq) significantly reduce genome size and complexity by only enriching RGA-like sequences and excluding the remaining genome [67]. RGASeq can efficiently provide high sequence coverage of the target regions only, without the need for high-coverage whole genome sequencing that would be required to produce a high-quality de novo genome assembly [68]. RGASeq relies on custom biotinylated RNA oligonucleotides, called baits or probes, which are complementary to RGA-like sequences. The target regions that hybridise to the baits are then 'fished' out and amplified. RenSeq performed on wheat NLRs produced an enriched library which contained a 1000-fold increase in NLRs compared to other genes [69]. Typically, tens of thousands of baits that show high sequence similarity to RGAs are used in any one reaction to effectively capture the full RGA repertoire. The effectiveness of RGASeq is largely influenced by the design of the baits, as they determine what part of the genome is targeted and enriched. Although only one gene model, ideally from the target species, is required to design a bait library, gene models from a wider range of sources, beyond that of the target species, are often incorporated into the design, for example Steuernagel et al. [69] included predicted NLRs from seven wheat relatives into their bait library design. The inclusion of gene models from closely related species expands the bait library to capture more RGA variation. As shown, many RGAs have been predicted in *Brassica* species, including in the pangenomes of *B. napus* and *B. oleracea* [5,45,51], and in over 30 genomes of various Brassicaceae species [50]. The NLR content of the well-studied *Brassica* relative, *A. thaliana*, has also been extensively catalogued [70]. These resources will help to create a diverse bait library designed from a wide range of closely related species that can be utilised for RGASeq in *Brassicas*.

Crop wild relatives (CWRs) are valuable sources of genetic diversity, as extensive breeding has rendered domesticated species genetically depauperate, meaning cultivated crops contain fewer RGAs and are more vulnerable to disease [71]. Therefore, it is important to explore the RGA diversity in CRWs and landraces to expand breeding resources [72]. The identification of novel RGAs in wild relatives, however, is difficult since most wild species lack reference genomes. Even when a reference genome is available, traditional approaches like GWAS are not able to detect RGAs that significantly differ to the reference, a problem compounded by the fact that *Brassica* RGAs were found to be highly polymorphic and show significant structural variation [5,51]. In response to this limitation, Arora et al. [62] devised an approach that combines association genetics and RenSeq (AgRenSeq), which analyses *k*-mers that differ between resistant and susceptible lines. *K*-mers are subsequences of DNA that are length *k* and can be used for alignment-free sequence analysis enabling association mapping in the absence of a reference genome [73]. Notably, the AgRenSeq approach led to the identification of four wheat stem rust NLR resistance genes, *Sr33*, *Sr45*, *Sr46* and *SrTA1662*, two of which are novel. First, a diverse panel of *Aegilops tauschii*, a wild progenitor of diploid bread wheat, was screened against six races of the wheat stem rust pathogen. The NLR complement of the accessions included in the panel was enriched using a bait library designed for *Ae. tauschii* and was filtered for NLR motifs. After association-based *k-mer* analysis, *k-mers* highly associated with resistance were directly overlaid onto NLR assemblies, revealing full-length stem rust resistance genes. A map-based and mutagenesis approach to identify *Sr46* was also conducted, producing a candidate which was identical to the one found through AgRenSeq, thereby confirming the accuracy and effectiveness of the newly developed AgRenSeq method.

A similar *k-mer* based association approach could be taken to pinpoint RLPs and RLKs, hence all RGA families could be captured through a combination of association genetics and targeted resistance gene enrichment, which we refer to as AgRGASeq. AgRGASeq can be harnessed to precisely identify causative RGAs in a diverse panel of *Brassica* germplasm in the absence of a reference genome. By utilising *k*-mers, this approach has the potential to capture the full diversity of RGAs, including those that greatly diverge from a reference and would otherwise go undetected. Its application in canola-blackleg resistance could be advantageous, as there are numerous closely related genera containing wild species which show resistance to blackleg, the majority of which have no or limited reference genomes [74–78]. Therefore, AgRGASeq could be used to unlock CWRs untapped RGA resources.

Steuernagel et al. [69] deployed a different approach to identify and clone causative NLRs in the large and complex hexaploid bread wheat genome, exploiting loss-of-resistance mutations caused by chemical mutagenesis and RenSeq (MutRenSeq). To test MutRenSeq, a previously cloned wheat stem rust NLR, *Sr33* [79], was identified through the enrichment and comparison of NLRs captured from six susceptible mutants with ethyl methanesulfonate (EMS) derived mutations for *Sr33* and a resistant *Sr33* wild-type. Causative mutations were found in the resulting NLR contigs and were identical to the previously discovered mutations. Steuernagel et al. then cloned two uncharacterised wheat stem rust NLRs, *Sr22* and *Sr45*, using a similar method. MutRenSeq offers a fast workflow to pinpoint functional RGAs without fine mapping or positional cloning efforts, given you have the appropriate resistant wild type harbouring only one qualitative resistance gene and loss of disease resistance mutants [80]. To date, no studies have conducted mutagenesis on *Brassicas* with the aim of increasing disease resistance, thus MutRenSeq is a promising modification of RenSeq that could open a new avenue for RGA identification within *Brassica* species.

One of the dominant long read sequencing platforms, Pacific BioSciences (PacBio) single molecule real time sequencing (SMRT) was used in conjunction with RenSeq (SMRT RenSeq) to identify and clone a late blight broad-spectrum NLR, *Rpi-amr3i*, from *Solanum americanum*, a wild diploid potato relative [81]. Using a mapping population, a bulked segregant analysis approach was coupled with RenSeq to produce an enriched NLR complement of one resistant and one susceptible accession followed by PacBio SMRT sequencing. The consensus long reads were then filtered for NLR-like sequences and six candidates were identified. *Rpi-amr3i* was confirmed as the causative NLR through

transgenic expression analysis. Here, SMRT RenSeq was able to obtain high coverage (5–70 X) long reads of potato NLRs, which captured the flanking promoter and terminator sequences, allowing for the assembly of the complete gene. The long reads offered by SMRT RenSeq may be key in overcoming the limitations posed by the large structural variation found in *Brassica* RGAs.

6. Building Pan-RGAomes to Capture Wider RGA Diversity

Due to the recognition that there is large genomic variation within a species, plant pangenomes are becoming widely adopted in plant genomics [38]. In a similar vein, the genome-wide annotation of RGAs from a single reference does not capture the presence/absence and SNP variation in the RGA repertoire within a species. Here, we reason that the pangenome concept can be applied to RGAs, where the RGA repertoire of many individuals of a species, together make up the 'pan-RGAome'. The pan-RGAome would capture the near-complete species-wide diversity of RGAs and include those that diverge from a single reference. Several methods would enable the construction of a pan-RGAome, including a whole genome resequencing (WGRS) approach, whereby RGAs would be predicted from a pangenome or through the use of RenSeq, which would involve targeting and enriching the RGA complement of many accessions of a species.

The near-complete species-wide NLR inventory, or pan-NLRome, in the model species *A. thaliana* was recently constructed using SMRT RenSeq and is the first example of defining a species-wide repertoire of plant immunity genes [70]. The NLR complement of 64 diverse *A. thaliana* accessions was enriched using a custom bait library, sequenced and reconstructed, producing 13,167 NLR gene annotations. The pan-NLRome was found to be almost completely defined with the NLRs from any 40 accessions out of the 64 included in the analysis, indicating that even though NLR loci show high variability, there is a finite number of divergent loci. In addition, the creation of the pan-NLRome enabled the identification of new NLR architectures, including some which contained integrated domains. The pan-NLRome for *A. thaliana* serves as a blueprint for future studies that wish to categorize species-wide RGA diversity. Importantly, the development of RLP/KSeq has allowed the capture of RGAs beyond the NLR family and through a combination of RenSeq and RLP/KSeq, building a pan-RGAome using targeted enrichment technology is now possible.

Although creating a pan-RGAome through RGA prediction in a pangenome is attainable, for example in *B. napus* and *B. oleracea* [5,51], it is limited to the accessions that are incorporated into the pangenome. In this regard, it is well-suited to species that already have reference genomes available for many accessions, usually well-studied major crops. If new lines of interest need to be added to the pan-RGAome, the reference genome of that line would also need to be incorporated into the pangenome before RGA prediction to ensure that any variation would map back to a reference. A RGASeq approach that utilises *k*-mer based analysis, on the other hand, like the one taken by Arora et al. [62], would enable the pan-RGAome in species lacking a reference genome to be defined, especially if used in conjunction with long-read sequencing. Characterising the pan-RGAome of wild *Brassica* relatives that demonstrate strong resistance towards disease would broaden the RGA pool available for improved cultivar breeding.

7. Conclusions

Brassica crops are threatened by diseases, such as blackleg, sclerotinia stem rot and clubroot. Understanding the genetic basis of disease resistance is crucial for managing these diseases. Construction of multiple *Brassica* genome and pangenome assemblies have greatly facilitated reference-based QTL mapping for disease resistance phenotypes. RGAs are found to be associated with the resistance mechanisms. Genome-wide and pangenome-wide prediction of RGAs and comparison of sequence variations of the RGAs located in previously identified QTL have contributed to efficient identification of RGA candidates. However, reference-based QTL mapping and cloning could be limited when exploring novel resistance in landraces and wild relatives which harbour greater genetic diversity.

Describing the full species-wide diversity of RGAs is an important initial step in identifying causative RGAs and critical in understanding the evolution of disease resistance genes. Building a pan-RGAome provides a comprehensive overview of RGA variation and architecture, which is highly relevant to *Brassica* species as there is significant RGA variation among individuals. It will also prove useful in narrowing down RGAs in previously described QTL by providing the near complete catalogue of potential candidate genes. Finally, defining the pan-RGAome will be essential in understanding the complex evolution of RGAs, which will provide insight into the mechanisms of disease resistance. The pan-RGAome is a promising and powerful resource which will help to improve cultivars through disease resistance breeding.

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

References

1. Lv, H.; Fang, Z.; Yang, L.; Zhang, Y.; Wang, Y. An update on the arsenal: Mining resistance genes for disease management of *Brassica* crops in the genomic era. *Hortic. Res.* **2020**, *7*, 34. [CrossRef] [PubMed]

2. Nagaharu, U. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn. J. Bot.* **1935**, *7*, 389–452.

3. Katche, E.; Quezada-Martinez, D.; Katche, E.I.; Vasquez-Teuber, P.; Mason, A.S. Interspecific hybridization for *Brassica* crop improvement. *Crop Breed. Genet. Genom.* **2019**, *1*, e190007.

4. Ghassemi-Golezani, K.; Hassanzadeh, N.; Shakiba, M.R.; Esmaeilpour, B. Exogenous salicylic acid and 24-epi-brassinolide improve antioxidant capacity and secondary metabolites of *Brassica nigra*. *Biocatal. Agric. Biotechnol.* **2020**, *26*, 101636. [CrossRef]

5. Bayer, P.E.; Golicz, A.A.; Tirnaz, S.; Chan, C.K.K.; Edwards, D.; Batley, J. Variation in abundance of predicted resistance genes in the *Brassica oleracea* pangenome. *Plant Biotechnol. J.* **2018**, *17*, 789–800. [CrossRef]

6. Van de Wouw, A.P.; Howlett, B.J. Advances in understanding the *Leptosphaeria maculans-Brassica* pathosystem and their impact on disease management. *Can. J. Plant Pathol.* **2019**, *42*, 149–163. [CrossRef]

7. Kayaçetİn, F. Performance of some important agronomic characteristics of *Brassica juncea* L. genotypes under fall sowing at two locations of Ankara, Turkey. *KSU J. Agric. Nat.* **2020**, *23*, 966–974.

8. Wang, T.; Guo, S.; Yu, K.; Zou, J.; Jiang, Y.; Wan, W.; Ye, B.; Yang, R.; Khattak, A.N.; Meng, J. iTRAQ-Based Comparative Proteomic Analysis in Mature Leaves of Ethiopian Mustard (*Brassica carinata*) Provides Insights into the Whole-Proteomic Profiles and the Role of Anthocyanin Biosynthesis in Leaf Color Diversity. 2019. Available online: https://www.researchsquare.com/article/rs-4986/v1 (accessed on 7 September 2020).

9. Fitt, B.D.L.; Brun, H.; Barbetti, M.J.; Rimmer, S.R. World-wide importance of phoma stem canker (*Leptosphaeria maculans* and *L. biglobosa*) on oilseed rape (*Brassica napus*). *Eur. J. Plant Pathol.* **2006**, *114*, 3–15. [CrossRef]

10. Van De Wouw, A.P.; Marcroft, S.J.; Howlett, B.J. Blackleg disease of canola in Australia. *Crop Pasture Sci.* **2016**, *67*, 273–283. [CrossRef]

11. Zhang, X.; Fernando, W.D. Insights into fighting against blackleg disease of *Brassica napus* in Canada. *Crop Pasture Sci.* **2018**, *69*, 40–47. [CrossRef]

12. Rouxel, T.; Penaud, A.; Pinochet, X.; Brun, H.; Gout, L.; Delourme, R.; Schmit, J.; Balesdent, M.H. A 10-year survey of populations of *Leptosphaeria maculans* in France indicates a rapid adaptation towards the *Rlm1* resistance gene of oilseed rape. *Eur. J. Plant Pathol.* **2003**, *109*, 871–881. [CrossRef]

13. Sprague, S.; Marcroft, S.; Hayden, H.; Howlett, B. Major gene resistance to blackleg in *Brassica napus* overcome within three years of commercial production in southeastern Australia. *Plant Dis.* **2006**, *90*, 190–198. [CrossRef] [PubMed]

14. Derbyshire, M.; Denton-Giles, M. The control of Sclerotinia stem rot on oilseed rape (*Brassica napus*): Current practices and future opportunities. *Plant Pathol.* **2016**, *65*, 859–877. [CrossRef]

15. Mei, J.; Qian, L.; Disi, J.; Yang, X.; Li, Q.; Li, J.; Frauen, M.; Cai, D.; Qian, W. Identification of resistant sources against *Sclerotinia sclerotiorum* in *Brassica* species with emphasis on *B. oleracea*. *Euphytica* **2011**, *177*, 393–399. [CrossRef]

16. Dixon, G.R. Clubroot (*Plasmodiophora brassicae* Woronin)–An agricultural and biological challenge worldwide. *Can. J. Plant Pathol.* **2014**, *36*, 5–18. [CrossRef]

17. Strelkov, S.; Tewari, J.; Smith-Degenhardt, E. Characterization of *Plasmodiophora brassicae* populations from Alberta, Canada. *Can. J. Plant Pathol.* **2006**, *28*, 467–474. [CrossRef]

18. Chai, A.; Xie, X.; Shi, Y.; Li, B. Research status of clubroot (*Plasmodiophora brassicae*) on cruciferous crops in China. *Can. J. Plant Pathol.* **2014**, *36*, 142–153. [CrossRef]

19. Nelson, R.; Wiesner-Hanks, T.; Wisser, R.; Balint-Kurti, P. Navigating complexity to breed disease-resistant crops. *Nat. Rev. Genet.* **2018**, *19*, 21–33. [CrossRef]

20. Araújo, A.C.D.; Fonseca, F.C.D.A.; Cotta, M.G.; Alves, G.S.C.; Miller, R.N.G. Plant NLR receptor proteins and their potential in the development of durable genetic resistance to biotic stresses. *Biotechnol. Res. Innov.* **2020**, *3*, 80–94.

21. Tang, D.; Wang, G.; Zhou, J.-M. Receptor kinases in plant-pathogen interactions: More than pattern recognition. *Plant Cell* **2017**, *29*, 618–637. [CrossRef]

22. Andersen, E.J.; Ali, S.; Byamukama, E.; Yen, Y.; Nepal, M.P. Disease resistance mechanisms in plants. *Genes* **2018**, *9*, 339. [CrossRef] [PubMed]

23. Zhou, Y.; Sun, L.; Wassan, G.M.; He, X.; Shaban, M.; Zhang, L.; Zhu, L.; Zhang, X. GbSOBIR1 confers *Verticillium* wilt resistance by phosphorylating the transcriptional factor GbbHLH171 in *Gossypium barbadense*. *Plant Biotechnol. J.* **2019**, *17*, 152–163. [CrossRef] [PubMed]

24. Li, W.; Deng, Y.; Ning, Y.; He, Z.; Wang, G.L. Exploiting broad-spectrum disease resistance in crops: From molecular dissection to breeding. *Annu. Rev. Plant Biol.* **2020**, *71*, 575–603. [CrossRef] [PubMed]

25. Tamborski, J.; Krasileva, K.V. Evolution of plant NLRs: From natural history to precise modifications. *Annu. Rev. Plant Biol.* **2020**, *71*, 355–378. [CrossRef] [PubMed]

26. Dardick, C.; Schwessinger, B.; Ronald, P. Non-arginine-aspartate (non-RD) kinases are associated with innate immune receptors that recognize conserved microbial signatures. *Curr. Opin. Plant Biol.* **2012**, *15*, 358–366. [CrossRef]

27. Wang, X.; Wang, H.; Wang, J.; Sun, R.; Wu, J.; Liu, S.; Bai, Y.; Mun, J.H.; Bancroft, I.; Cheng, F.; et al. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* **2011**, *43*, 1035–1039. [CrossRef]

28. Belser, C.; Istace, B.; Denis, E.; Dubarry, M.; Baurens, F.C.; Falentin, C.; Genete, M.; Berrabah, W.; Chèvre, A.M.; Delourme, R.; et al. Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nat. Plants* **2018**, *4*, 879–887. [CrossRef]

29. Liu, S.; Liu, Y.; Yang, X.; Tong, C.; Edwards, D.; Parkin, I.A.; Zhao, M.; Ma, J.; Yu, J.; Huang, S.; et al. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* **2014**, *5*, 3930. [CrossRef]

30. Parkin, I.A.P.; Koh, C.S.; Tang, H.; Robinson, S.J.; Kagale, S.; Clarke, W.E.; Town, C.D.; Nixon, J.; Krishnakumar, V.; Bidwell, S.L.; et al. Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biol.* **2014**, *15*, R77. [CrossRef]

31. Sun, D.; Wang, C.; Zhang, X.; Zhang, W.; Jiang, H.; Yao, X.; Liu, L.; Wen, Z.; Niu, G.; Shan, X. Draft genome sequence of cauliflower (*Brassica oleracea* L. var. *botrytis*) provides new insights into the C genome in Brassica species. *Hortic. Res.* **2019**, *6*, 82.

32. Chalhoub, B.; Denoeud, F.; Liu, S.; Parkin, I.A.P.; Tang, H.; Wang, X.; Chiquet, J.; Belcram, H.; Tong, C.; Samans, B.; et al. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* **2014**, *345*, 950–953. [CrossRef] [PubMed]

33. Bayer, P.E.; Hurgobin, B.; Golicz, A.A.; Chan, C.K.K.; Yuan, Y.; Lee, H.; Renton, M.; Meng, J.; Li, R.; Long, Y.; et al. Assembly and comparison of two closely related *Brassica napus* genomes. *Plant Biotechnol. J.* **2017**, *15*, 1602–1610. [CrossRef] [PubMed]

34. Sun, F.; Fan, G.; Hu, Q.; Zhou, Y.; Guan, M.; Tong, C.; Li, J.; Du, D.; Qi, C.; Jiang, L.; et al. The high-quality genome of *Brassica napus* cultivar 'ZS11' reveals the introgression history in semi-winter morphotype. *Plant J.* **2017**, *92*, 452–468. [CrossRef] [PubMed]

35. Song, J.M.; Guan, Z.; Hu, J.; Guo, C.; Yang, Z.; Wang, S.; Liu, D.; Wang, B.; Lu, S.; Zhou, R.; et al. Eight high-quality genomes reveal pangenome architecture and ecotype differentiation of *Brassica napus*. *Nat. Plants* **2020**, *6*, 34–45. [CrossRef] [PubMed]

36. Paritosh, K.; Yadava, S.K.; Singh, P.; Bhayana, L.; Mukhopadhyay, A.; Gupta, V.; Bisht, N.C.; Zhang, J.; Kudrna, D.A.; Copetti, D.; et al. A chromosome-scale assembly of allotetraploid *Brassica juncea* (AABB) elucidates comparative architecture of the A and B genomes. *Plant Biotechnol. J.* **2020**. [CrossRef] [PubMed]

37. Yang, J.; Liu, D.; Wang, X.; Ji, C.; Cheng, F.; Liu, B.; Hu, Z.; Chen, S.; Pental, D.; Ju, Y.; et al. The genome sequence of allopolyploid *Brassica juncea* and analysis of differential homoeolog gene expression influencing selection. *Nat. Genet.* **2016**, *48*, 1225–1232. [CrossRef] [PubMed]

38. Bayer, P.E.; Golicz, A.A.; Scheben, A.; Batley, J.; Edwards, D. Plant pangenomes are the new reference. *Nat. Plants* **2020**, *6*, 914–920. [CrossRef]

39. Golicz, A.A.; Bayer, P.E.; Barker, G.C.; Edger, P.P.; Kim, H.; Martinez, P.A.; Chan, C.K.K.; Severn-Ellis, A.; McCombie, W.R.; Parkin, I.A.P.; et al. The pangenome of an agronomically important crop plant *Brassica oleracea*. *Nat. Commun.* **2016**, *7*, 13390. [CrossRef]

40. Hurgobin, B.; Golicz, A.A.; Bayer, P.E.; Chan, C.K.K.; Tirnaz, S.; Dolatabadian, A.; Schiessl, S.V.; Samans, B.; Montenegro, J.D.; Parkin, I.A. Homoeologous exchange is a major cause of gene presence/absence variation in the amphidiploid *Brassica napus*. *Plant Biotechnol. J.* **2018**, *16*, 1265–1274. [CrossRef]

41. Li, P.; Quan, X.; Jia, G.; Xiao, J.; Cloutier, S.; You, F.M. RGAugury: A pipeline for genome-wide prediction of resistance gene analogs (RGAs) in plants. *BMC Genom.* **2016**, *17*, 852. [CrossRef]

42. Mun, J.H.; Yu, H.J.; Park, S.; Park, B.S. Genome-wide identification of NBS-encoding resistance genes in *Brassica rapa*. *Mol. Genet. Genom.* **2009**, *282*, 617. [CrossRef] [PubMed]

43. Yu, J.; Tehrim, S.; Zhang, F.; Tong, C.; Huang, J.; Cheng, X.; Dong, C.; Zhou, Y.; Qin, R.; Hua, W.; et al. Genome-wide comparative analysis of NBS-encoding genes between *Brassica* species and *Arabidopsis thaliana*. *BMC Genom.* **2014**, *15*, 3. [CrossRef] [PubMed]

44. Zhang, Y.M.; Shao, Z.Q.; Wang, Q.; Hang, Y.Y.; Xue, J.Y.; Wang, B.; Chen, J.Q. Uncovering the dynamic evolution of nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes in Brassicaceae. *J. Integr. Plant Biol.* **2016**, *58*, 165–177. [CrossRef] [PubMed]

45. Alamery, S.; Tirnaz, S.; Bayer, P.; Tollenaere, R.; Chaloub, B.; Edwards, D.; Batley, J. Genome-wide identification and comparative analysis of NBS-LRR resistance genes in *Brassica napus*. *Crop Pasture Sci.* **2018**, *69*, 72–93. [CrossRef]

46. Fu, Y.; Zhang, Y.; Mason, A.S.; Lin, B.; Zhang, D.; Yu, H.; Fu, D. NBS-encoding genes in *Brassica napus* evolved rapidly after allopolyploidization and co-localize with known disease resistance loci. *Front. Plant Sci.* **2019**, *10*, 26. [CrossRef]

47. Rameneni, J.J.; Lee, Y.; Dhandapani, V.; Yu, X.; Choi, S.R.; Oh, M.H.; Lim, Y.P. Genomic and post-translational modification analysis of leucine-rich-repeat receptor-like kinases in *Brassica rapa*. *PLoS ONE* **2015**, *10*, e0142255. [CrossRef]

48. Liu, Z.; Xie, J.; Wang, H.; Zhong, X.; Li, H.; Yu, J.; Kang, J. Identification and expression profiling analysis of NBS–LRR genes involved in *Fusarium oxysporum* f. sp. *conglutinans* resistance in cabbage. *3 Biotech* **2019**, *9*, 202.

49. Stotz, H.U.; Harvey, P.J.; Haddadi, P.; Mashanova, A.; Kukol, A.; Larkan, N.J.; Borhan, M.H.; Fitt, B.D.; Raman, H. Genomic evidence for genes encoding leucine-rich repeat receptors linked to resistance against the eukaryotic extra-and intracellular *Brassica napus* pathogens *Leptosphaeria maculans* and *Plasmodiophora brassicae*. *PLoS ONE* **2018**, *13*, e0198201. [CrossRef]

50. Tirnaz, S.; Bayer, P.; Inturrisi, F.; Zhang, F.; Yang, H.; Dolatabadian, A.; Neik, T.X.; Severn-Ellis, A.; Patel, D.; Ibrahim, M.I.; et al. Resistance gene analogs in the Brassicaceae: Identification, characterization, distribution, and evolution. *Plant Physiol.* **2020**, *184*, 909–922. [CrossRef]

51. Dolatabadian, A.; Bayer, P.E.; Tirnaz, S.; Hurgobin, B.; Edwards, D.; Batley, J. Characterization of disease resistance genes in the *Brassica napus* pangenome reveals significant structural variation. *Plant Biotechnol. J.* **2020**, *18*, 969–982. [CrossRef]

52. Inturrisi, F.C. Genome-Wide Analysis of NBS-LRR Genes in Indian Mustard (*Brassica juncea*) and Prediction of Candidate Disease Resistance Genes. Ph.D. Thesis, University of Western Australia, Perth, Australia, December 2018.

53. Inturrisi, F.; Bayer, P.E.; Yang, H.; Tirnaz, S.; Edwards, D.; Batley, J. Genome-wide identification and comparative analysis of resistance genes in *Brassica juncea*. *Mol. Breed.* **2020**, *40*, 1–14. [CrossRef]

54. Bayer, P.E.; Edwards, D.; Batley, J. Bias in resistance gene prediction due to repeat masking. *Nat. Plants* **2018**, *4*, 762–765. [CrossRef] [PubMed]

55. Zhang, Y.; Neik, T.X.; Amas, J.; Cantila, A.; Saad, N.S.M.; Wu, T.; Batley, J. DNA-based screening of *Brassica* germplasms. In *Plant Genetic Resources: A Review of Current Research and Future Needs*; Dulloo, E., Ed.; Burleigh Dodds Science Publishing: Sawston, Cambridge, UK, 2020, in press.

56. Qasim, M.U.; Zhao, Q.; Shahid, M.; Samad, R.A.; Ahmar, S.; Wu, J.; Fan, C.; Zhou, Y. Identification of QTLs containing resistance genes for Sclerotinia Stem Rot in *Brassica napus* using comparative transcriptomic studies. *Front. Plant Sci.* **2020**, *11*, 776. [CrossRef] [PubMed]

57. Laila, R.; Park, J.I.; Robin, A.H.K.; Natarajan, S.; Vijayakumar, H.; Shirasawa, K.; Isobe, S.; Kim, H.T.; Nou, I.S. Mapping of a novel clubroot resistance QTL using ddRAD-seq in Chinese cabbage (*Brassica rapa* L.). *BMC Plant Biol.* **2019**, *19*, 13. [CrossRef]

58. Gabur, I.; Chawla, H.S.; Lopisso, D.T.; von Tiedemann, A.; Snowdon, R.J.; Obermeier, C. Gene presence-absence variation associates with quantitative *Verticillium longisporum* disease resistance in *Brassica napus*. *Sci. Rep.* **2020**, *10*, 4131. [CrossRef]

59. Hatakeyama, K.; Suwabe, K.; Tomita, R.N.; Kato, T.; Nunome, T.; Fukuoka, H.; Matsumoto, S.; Wu, K. Identification and characterization of *Crr1a*, a gene for resistance to clubroot disease (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. *PLoS ONE* **2013**, *8*, e54745. [CrossRef]

60. Jin, M.; Lee, S.S.; Ke, L.; Kim, J.S.; Seo, M.S.; Sohn, S.H.; Park, B.S.; Bonnema, G. Identification and mapping of a novel dominant resistance gene, *TuRB07* to Turnip mosaic virus in *Brassica rapa*. *Theor. Appl. Genet.* **2014**, *127*, 509–519. [CrossRef]

61. Larkan, N.J.; Ma, L.; Borhan, M.H. The *Brassica napus* receptor-like protein RLM2 is encoded by a second allele of the *LepR3/Rlm2* blackleg resistance locus. *Plant Biotechnol. J.* **2015**, *13*, 983–992. [CrossRef]

62. Arora, S.; Steuernagel, B.; Gaurav, K.; Chandramohan, S.; Long, Y.; Matny, O.; Johnson, R.; Enk, J.; Periyannan, S.; Singh, N.; et al. Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nat. Biotechnol.* **2019**, *37*, 139–143. [CrossRef]

63. Broekema, R.; Bakker, O.; Jonkers, I. A practical view of fine-mapping and gene prioritization in the post-genome-wide association era. *Open Biol.* **2020**, *10*, 190221. [CrossRef]

64. Liu, H.J.; Yan, J. Crop genome-wide association study: A harvest of biological relevance. *Plant J.* **2019**, *97*, 8–18. [CrossRef] [PubMed]

65. Jupe, F.; Witek, K.; Verweij, W.; Śliwka, J.; Pritchard, L.; Etherington, G.J.; Maclean, D.; Cock, P.J.; Leggett, R.M.; Bryan, G.J. Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *Plant J.* **2013**, *76*, 530–544. [CrossRef] [PubMed]

66. Lin, X.; Armstrong, M.; Baker, K.; Wouters, D.; Visser, R.G.; Wolters, P.J.; Hein, I.; Vleeshouwers, V.G. RLP/K enrichment sequencing; a novel method to identify receptor-like protein (RLP) and receptor-like kinase (RLK) genes. *New Phytol.* **2020**, *227*, 1264–1276. [CrossRef] [PubMed]

67. Hodges, E.; Xuan, Z.; Balija, V.; Kramer, M.; Molla, M.N.; Smith, S.W.; Middle, C.M.; Rodesch, M.J.; Albert, T.J.; Hannon, G.J. Genome-wide *in situ* exon capture for selective resequencing. *Nat. Genet.* **2007**, *39*, 1522–1527. [CrossRef]

68. Parla, J.S.; Iossifov, I.; Grabill, I.; Spector, M.S.; Kramer, M.; McCombie, W.R. A comparative analysis of exome capture. *Genome Biol.* **2011**, *12*, R97. [CrossRef]

69. Steuernagel, B.; Periyannan, S.K.; Hernández-Pinzón, I.; Witek, K.; Rouse, M.N.; Yu, G.; Hatta, A.; Ayliffe, M.; Bariana, H.; Jones, J.D. Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* **2016**, *34*, 652–655. [CrossRef]

70. Van de Weyer, A.L.; Monteiro, F.; Furzer, O.J.; Nishimura, M.T.; Cevik, V.; Witek, K.; Jones, J.D.; Dangl, J.L.; Weigel, D.; Bemm, F. A species-wide inventory of NLR genes and alleles in *Arabidopsis thaliana*. *Cell* **2019**, *178*, 1260–1272. [CrossRef]

71. Smýkal, P.; Nelson, M.N.; Berger, J.D.; Von Wettberg, E.J. The impact of genetic changes during crop domestication. *Agronomy* **2018**, *8*, 119. [CrossRef]

72. Zhang, F.; Batley, J. Exploring the application of wild species for crop improvement in a changing climate. *Curr. Opin. Plant Biol.* **2020**, *56*, 218–222. [CrossRef]

73. Rahman, A.; Hallgrímsdóttir, I.; Eisen, M.; Pachter, L. Association mapping from sequencing reads using k-mers. *Elife* **2018**, *7*, e32920. [CrossRef]

74. Gugel, R.; Seguin-Swartz, G.; Petrie, G. Pathogenicity of three isolates of *Leptosphaeria maculans* on *Brassica* species and other crucifers. *Can. J. Plant Pathol.* **1990**, *12*, 75–82. [CrossRef]

75. Tewari, J.; Bansal, V.; Tewari, I.; Stringam, G.; Thiagarajah, M. Reactions of some wild and cultivated accessions of *Eruca* against *Leptosphaeria maculans*. *Crucif. Newsl.* **1996**, *18*, 130–1331.

76. Chen, C.; Séguin-Swartz, G. Reaction of wild crucifers to *Leptosphaeria maculans*, the causal agent of blackleg of crucifers. *Can. J. Plant Pathol.* **1999**, *21*, 361–367. [CrossRef]

77. Snowdon, R.; Winter, H.; Diestel, A.; Sacristan, M. Development and characterisation of *Brassica napus-Sinapis arvensis* addition lines exhibiting resistance to *Leptosphaeria maculans*. *Theor. Appl. Genet.* **2000**, *101*, 1008–1014. [CrossRef]

78. Purnamasari, M.I.; Erskine, W.; Croser, J.S.; You, M.P.; Barbetti, M.J. Comparative reaction of *Camelina sativa* to *Sclerotinia sclerotiorum* and *Leptosphaeria maculans*. *Plant Dis.* **2019**, *103*, 2884–2892. [CrossRef]

79. Periyannan, S.; Moore, J.; Ayliffe, M.; Bansal, U.; Wang, X.; Huang, L.; Deal, K.; Luo, M.; Kong, X.; Bariana, H. The gene *Sr33*, an ortholog of barley *Mla* genes, encodes resistance to wheat stem rust race Ug99. *Science* **2013**, *341*, 786–788. [CrossRef]

80. Sahu, P.K.; Sao, R.; Mondal, S.; Vishwakarma, G.; Gupta, S.K.; Kumar, V.; Singh, S.; Sharma, D.; Das, B.K. Next generation sequencing based forward genetic approaches for identification and mapping of causal mutations in crop plants: A comprehensive review. *Plants* **2020**, *9*, 1355. [CrossRef]

81. Witek, K.; Jupe, F.; Witek, A.I.; Baker, D.; Clark, M.D.; Jones, J.D. Accelerated cloning of a potato late blight–resistance gene using RenSeq and SMRT sequencing. *Nat. Biotechnol.* **2016**, *34*, 656–660. [CrossRef]