Genome Biology

RESEARCH

Open Access

Allelic structural variation at the *NLR25-1* locus enhances defense against *Pseudomonas syringae* in kiwifruit

Ying Wu[1†], Yingzhen Wang[2†], Yunzhi Lin[1†], Meng Zhao[1], Yuyu Huo[1], Tao Zhang[1], Hongtao Wang[1], Yanyan Zhu[1], Lihuan Wang[1], Pengpeng Zheng[1], Junyang Yue[1*], Songhu Wang[1*] and Yongsheng Liu[1*]

[†]Ying Wu, Yingzhen Wang and Yunzhi Lin contributed equally to this work.

*Correspondence: yuejy@ahau.edu.cn; wangsonghu@ahau.edu.cn; liuyongsheng1122@ahau.edu.cn

[1] Anhui Province Key Laboratory of Horticultural Crop Quality Biology, School of Horticulture, Anhui Agricultural University, Hefei 230036, China
[2] School of Forestry Science and Technology, Lishui Vocational and Technical College, Lishui 323000, China

Abstract

Background: Canker disease caused by *Pseudomonas syringae* pv. *actinidiae* (*Psa*) poses a major threat to cultivated kiwifruit, and utilization of wild relatives are key to improve resistance. However, comprehensive comparative genomic analyses between cultivated kiwifruit and their wild relatives with enhanced resistance to Psa remain limited.

Results: Here we generate chromosome-scale genome assemblies for eleven wild *Actinidia eriantha* accessions and one interspecific hybrid between *Actinidia eriantha* and cultivated *Actinidia chinensis* var. *chinensis*. Integrating these with twelve previously released genomes including three *Actinidia eriantha* and nine *Actinidia chinensis* var. *chinensis*, we construct a reference-unbiased graph-based pangenome. These datasets reveal extensive genomic variation, including 31,790,044 SNPs, 13,512,079 InDels and 623,478 structural variations, and provide a landscape of structural variations within and between the two species. Leveraging these datasets, we identify a wild allele showing allele-specific expression, *AeNLR25-1*, which enhances *Psa* resistance in cultivated kiwifruit. Genetic and molecular analyses demonstrate that a transposable element-induced structural variation in the *AeNLR25-1* promoter introduces a species-specific WRKY binding site, conferring enhanced defense against *Psa*.

Conclusions: Pangenome across cultivated species and wild relatives provides a theoretical framework for accelerating kiwifruit genetic improvement through pangenome-enabled identification of favorable wild alleles.

Keywords: *Actinidia*, Pangenome, Structural variation (SV), Allele-specific expression (ASE), *Pseudomonas syringae* pv. *actinidiae (Psa)*

Background

The successful domestication of plants and animals was one of the most important events in the history of human civilization. As one of the most recently domesticated fruit crops, kiwifruit (*Actinidia* ssp.) has gained popularity worldwide due to its high

vitamin C content, balanced nutritional composition of minerals, dietary fiber, and other beneficial metabolic products. Native to China, kiwifruit seed was initially introduced to New Zealand in 1904 and since then the domestication of kiwifruit began [1]. In recent years, kiwifruit cultivation has rapidly expanded to the world's temperate regions and the crop currently covers approximately 286,348 hectares and produces nearly 4.433 million tons of fruit in 2023 (http://faostat.fao.org).

However, the biggest challenge for the cultivated kiwifruit (*A. chinensis* var. *chinensis*/*A. chinensis* var. *deliciosa* complex, with diploid, tetraploid and hexaploid varieties) is the outbreak of canker disease caused by *Pseudomonas syringae* pv. *actinidiae* (*Psa*) [2, 3]. *Psa* was first reported in Japan as the causal agent of the bacterial canker of kiwifruit [4]. Subsequently, this bacterial pathogen was recorded and prevalent in all major kiwifruit-producing countries [5–9], resulting in substantial economic losses [10, 11]. The cultivated kiwifruit lacks genetic diversity particularly for *Psa* resistance-related haplotypes, possibly due to intensive artificial selections in fruit yield and quality.

By contrast, wild relatives of *Actinidia* have adapted to diverse ecological environments and manifest broad genetic and phenotypic diversity [12]. These wild species are characterized with a rich source of allelic variation and contain genes for biotic and abiotic stress tolerance, as well as consumer preferred traits, such as colorful flesh, high levels of soluble solids content, flavor compounds, vitamin C content, and other nutritional metabolites [13–18]. Therefore, effectively utilizing the natural diversity of these wild germplasms is crucial for promoting genetic improvement of kiwifruit.

The assembly of several haplotype-resolved *Actinidia* reference genomes has allowed comprehensive characterization of genetic diversity in terms of SNPs and small insertion/deletions (indels) between individual haplotypes, unraveling favorable trait domestication and adaptive evolution [19–23]. Nonetheless, accumulating evidence have implicated that large structural variants (SVs), such as copy number variants (CNVs) and presence/absence variants (PAVs), also play pivotal roles in plant functional diversity and adaptive evolution [24]. Recent advances in kiwifruit genomics include a pangenome constructed by using 14 chromosome-scale haplotype-resolved genome assemblies from seven diploid representative accessions in *A. chinensis* var. *chinensis* [20], and a super-pangenome build from 15 high-quality assemblies of eight *Actinidia* species [25]. These analyses demonstrated that SVs contribute to phenotypic variance/diversity and can be utilized in functional genomic studies and accelerating breeding design.

Indeed, the genus *Actinidia* includes 54 species and 75 taxa and exhibits tremendous genetic variation/diversity within or between *Actinidia* species in its botanical or horticultural characteristics, such as organ morphology, bud germination time, flowering time, fruit yield/quality and adaptability to various environments [26–28]. Except for the cultivated *A. chinensis* var. *chinensis*/*A. chinensis* var. *deliciosa*, the vast majority of wild relatives are unexplored. Among them, *A. eriantha* is of particular interest to breeders due to its distinct resistance to *Psa*, exceptionally high vitamin C content, relatively short juvenile phase and interspecific crossability [29]. This species is widely distributed in the large area of the middle and lower reaches in Yangtze River Basin of China, and has been subjected to breeding selection and interspecific hybridization and introgression [12, 30–32]. Currently, only a limited number of genome assemblies for *A. eriantha* are available, which hinders the dissection and utilization of its genetic variants [14, 31, 33,

34]. To fully explore and utilize the favorable genetic resources of *A. eriantha*, it is necessary to assemble additional genomes to facilitate effective SVs detection and their uses in genetic improvement of kiwifruit.

In this study, combining with the previously released chromosome-level reference genomes from 12 cultivated *A. chinensis* var. *chinensis* accessions [20, 22, 23, 35], we constructed a reference-unbiased graph-based pangenome by de novo assembling chromosome-level genomes of an additional eleven wild *A. eriantha* accessions collected from its natural distribution area and one artificial hybrid individual between *A. eriantha* and *A. chinensis* var. *chinensis*. Comparative analyses reveal the distinct structural variations across the tested species, enabling the discovery of a wild haplotype harboring a highly conserved WRKY-binding *cis*-element in an SV embedded in the promoter of an NLR gene that has the potential to increase resistance against *Psa* in cultivated kiwifruit. Our findings provide insights into the exploitation of wild genetic resources by constructing interspecific pangenome across diverse species for crop improvement.

Results

Chromosome-level genome *de novo* assemblies and annotation of 12 kiwifruit accessions

To capture the genomic variations, 11 wild *A. eriantha* accessions (LD19 and YX01 are male, and the rest are female) representing its natural distribution areas and an interspecific hybrid F1 individual (named 'HH-1') between an *A. eriantha* accession 'White' [33] and an *A. chinensis* var. *chinensis* accession 'H0809' [35] were selected for de novo genome assembly (Table 1). All the tested accessions are diploid and their fruit characteristics are highly variable (Additional file 1: Table S1, Additional file 2: Fig. S1). Using the PacBio Revio sequencing platform, a total of 14.1 to 64.4 Gb of HiFi reads (approximately 21.7 to 92.0 ×) were generated (Additional file 1: Table S2) and then these reads were assembled to contigs of two haplotypes (primary haplotype and alternate haplotype) using hifiasm (v0.16.1) [36]. For the 11 wild *A. eriantha* accessions, contigs were anchored onto 29 chromosomes using the recently released telomere-to-telomere (T2T) genomes of MDHAPA (*A. eriantha*) [31]. For 'HH-1', contigs were assigned to two chromosome-scale haplotypes (haplotype 1 sourced from *A. eriantha*, here after HH-1.hap1; and haplotype 2 sourced from *A. chinensis* var. *chinensis,* here after HH-1. hap2) using quarTeT software [37] based on their collinearities with the T2T reference genomes [31, 35]. Finally, a total of 24 de novo haplotype assemblies were generated and the final assembly sizes ranged from 581.2 to 630.7 Mb with contig N50 ranged from 5 to 21 Mb (Table 1). The BUSCO completeness of these assemblies was estimated to be 98.62% on average, with the lowest score 94.00% (Table 1). All the long terminal repeat (LTR) assembly index (LAI) scores exceed 18.94 and thus meet the "reference" level [38], ensuring the quality of the 24 haplotype assemblies (Table 1). Next, we utilized the EDTA toolkit [39] to annotate transposable elements (TE) and reveal the repetitive sequences' contents ranging from 38.14% to 41.92% (Table 1, Additional file 1: Table S3). Combining the previously published genomes (Additional file 1: Table S4) [22, 23, 31, 35, 40, 41], we found that the average genome size of *A. eriantha* group was approximately 12.5 Mb larger than that of *A. chinensis* var. *chinensis* group (Additional file 2: Fig. S2a, Additional file 1: Table S5). Additionally, the average TE content in *A. eriantha* (40.9%) was slightly higher than in *A. chinensis* var. *chinensis* (38.6%) (Additional file 2: Fig. S2b, Additional

Table 1 Statistics of the assembly and annotation of 12 newly sequenced kiwifruit genomes in this study

Sample	Species	City	Location	Accession	Total size of contigs (Mb)	Number of contigs	N50 (Mb)	Total size of genomes (Mb)	gaps	TE size (%)	Number of genes	Genome BUSCOs (%)	LAI
CY38	*A. eriantha*	Ganzhou City, Jiangxi Province, China	25.549876 N, 114.307938 E	CY38.hap1	646	460	19	629.4	23	41.92	47,199	99.20	20.43
				CY38.hap2	629	216	17	621.1	31	40.95	45,477	98.60	20.19
DG01	*A. eriantha*	Qiandongnan City, Guizhou Province, China	26.149891 N, 109.051666 E	DG01.hap1	772	3,777	13	613.1	75	40.72	45,108	98.60	20.17
				DG01hap2	688	1,257	14	620.4	47	40.85	46,552	99.00	20.56
GZ26	*A. eriantha*	Nanping City, Fujian Province, China	27.521293 N, 117.260915 E	GZ26.hap1	660	707	21	630.7	6	40.64	46,457	98.90	21.04
				GZ26.hap2	617	211	19	607.5	21	40.36	45,294	99.00	20.32
HN01	*A. eriantha*	Zhuzhou City, Hunan Province, China	27.068227 N, 113.5619189 E	HN01.hap1	639	376	20	623.2	21	41.42	45,872	99.40	21.90
				HN01.hap2	626	226	18	616.9	42	40.93	45,907	98.70	21.52
LD19	*A. eriantha*	Lishui City, Zhejiang Province, China	28.045319 N, 120.064087 E	LD19.hap1	622	573	5	603.5	194	40.66	44,858	94.00	20.67
				LD19.hap2	627	343	10	581.2	75	41.38	42,742	97.70	21.34
MQ32	*A. eriantha*	Fuzhou City, Fujian Province, China	26.053724 N, 118.882756 E	MQ32.hap1	640	590	20	619.9	24	40.96	46,402	99.20	21.11
				MQ32.hap2	627	344	17	612.6	33	40.89	45,467	99.00	20.73
QS41	*A. eriantha*	Shangrao City, Jiangxi Province, China	28.297126 N, 117.745972 E	QS41.hap1	643	573	21	622.6	17	40.74	45,910	99.30	21.36
				QS41.hap2	625	286	20	612.8	11	40.99	45,301	98.90	20.79
RY10	*A. eriantha*	Shaoguan City, Guangdong Province, China	24.674474 N, 113.056183 E	RY10.hap1	756	3,884	11	615.9	54	41.00	45,543	99.10	21.15
				RY10.hap2	676	1,367	14	593.2	43	40.79	43,704	97.30	20.78
YL01	*A. eriantha*	Wenzhou City, Zhejiang Province, China	27.518015 N, 119.814148 E	YL01.hap1	642	488	21	623.3	5	41.29	46,553	99.00	21.82
				YL01.hap2	613	262	19	602.9	29	40.47	45,423	98.60	20.82
YX01	*A. eriantha*	Sanming City, Fujian Province, China	26.165815 N, 118.197579 E	YX01.hap1	639	485	21	619.1	26	41.30	45,596	98.50	21.31
				YX01.hap2	624	198	19	615.3	22	40.74	45,589	99.20	20.33

Table 1 (continued)

Sample	Species	City	Location	Accession	Total size of contigs (Mb)	Number of contigs	N50 (Mb)	Total size of genomes (Mb)	gaps	TE size (%)	Number of genes	Genome BUSCOs (%)	LAI
YX02	*A. eriantha*	Sanming City, Fujian Province, China	26.156546 N, 118.192313 E	YX02.hap1	638	370	21	621.7	6	41.03	45,923	99.00	21.19
				YX02.hap2	615	225	20	604.2	8	40.79	44,667	98.60	20.94
HH-1	interspecific hybrid between *A. chinensis* var. *chinensis and A. eriantha*			HH-1.hap1	655	751	20.7	612.5	2	40.10	48,555	98.90	24.19
				HH-1.hap2	612	278	20.8	600.9	2	38.14	48,034	99.10	18.94

file 1: Table S3). Obviously, the increased genome size can be primarily explained by the corresponding increase of TE size (Additional file 2: Fig. S2c). By integrating transcriptome-based, ab initio, and homologous protein-based prediction methods, a total of 42,742 to 48,555 genes across the 24 assemblies were finally predicted (Table 1).

Gene family-based pangenome architecture

First, we constructed a pangenome of *A. eriantha* by utilizing 15 accessions from different geographical locations with 28 haplotypes/assemblies, including the 23 present de novo haplotypes and 5 previously published haplotypes/assemblies (Additional file 1: Table S4) [31, 35, 41]. Orthologs investigation allowed classification of annotated genes from the 28 assemblies into 54,770 gene families (Additional file 2: Fig. S3a). The total number of gene families increased gradually by adding an additional assembly, and approached a plateau at $n=15$ (Additional file 2: Fig. S3a), indicating the pangenome reached a saturated state. Among these 54,770 gene families, 14,071 gene families (25.7%) were defined as core gene families (present in all 28 assemblies), 8,121 gene families (14.8%) as softcore families (present in 26 to 27 assemblies), 32,416 gene families (59.2%) as dispensable gene families (present in 2 to 25 assemblies), and 162 gene families (0.3%) as private gene families (present in only a single assembly) (Additional file 2: Fig. S3b).

Since an interspecific pangenome construction allows integrating the entire genetic variation across cultivated species and wild relatives [42], we subsequently add 18 *A. chinensis* var. *chinensis* haplotypes/assemblies (including hap2 of 'HH-1' and 17 previously published haplotypes/assemblies from 10 diploid accessions) (Additional file 1: Table S4) into the constructed pangenome of *A. eriantha*, creating an actually interspecific pangenome across both species [20, 22, 23, 31, 35, 41]. Phylogenetic analysis revealed that the sampled assemblies of *A. eriantha* and *A. chinensis* var. *chinensis* are clustered into distinct groups and those haplotypic assemblies derived from the F1 hybrid are separately clustered into their corresponding parental groups (Additional file 2: Fig. S4). All the annotated genes from the 24 accessions (i.e., 14 *A. eriantha,* nine *A. chinensis* var. *chinensis* and one hybrid) with 46 assemblies were classified into 56,572 gene families by orthologs analysis (Fig. 1a), slightly more than those in the pangenome of *A. eriantha* alone. The saturation curve analysis showed that the number of gene families ceased to increase while the number of assemblies approached 30, indicating that the sample size has reached or is beyond saturation (Fig. 1a). After incorporating *A. chinensis* var. *chinensis* genomes, we observed a reduction in core gene families to 21.4% and softcore gene families to 11.1%, alongside an increase in dispensable gene families to 67.2% and a decrease in private gene families to 0.3% (Fig. 1b). Meanwhile, we found that the exon count, CDS length, and gene expression levels of core gene families were higher than those of non-core genes (Additional file 2: Fig. S5), indicating the crucial role of core genes in plant growth and development. GO enrichment analysis revealed that core gene families are primarily involved in fundamental biological processes, such as cell wall biogenesis, anion transmembrane transport, shoot system morphogenesis, and vegetative to reproductive phase transition of meristem (Additional file 2: Fig.

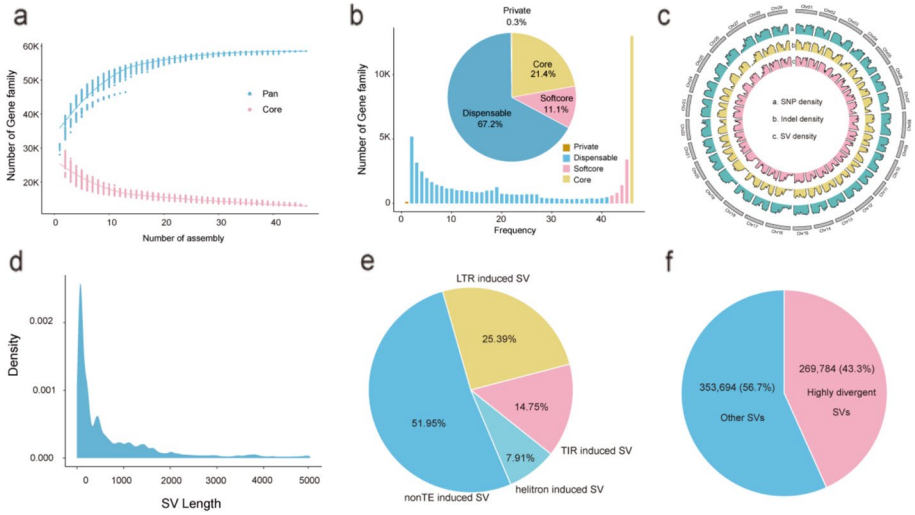

Fig. 1 Construction and analysis of an interspecific pangenome for *A. eriantha* and *A. chinensis* var. *chinensis* accessions. **a** Variation of gene families in the pan-genome and core-genome with the increase in number of genome assemblies. **b** Compositions of the interspecific pan-genome and individual genomes. The histogram shows the number of gene families in the 46 assemblies with different frequencies. The pie chart shows the proportion of the gene family marked by each composition. **c** SV density across all 22 accessions. **d** Length distribution of SVs in the graph pangenome. **e** The ratio of TE-induced SVs. **f** The ratio of highly divergent SVs between *A. chinensis* var. *chinensis* and *A. eriantha* accessions

S6a), whereas the variable gene families are enriched in biological processes in terms of immune response and/or secondary metabolism (Additional file 2: Fig. S6b).

Structural variations of the assembled genomes

A reference-unbiased pangenome is crucial for identifying structural variations (SVs) and capturing the genetic diversity distributed in different accessions. Thus, we constructed a reference-unbiased pangenome using the PGGB pipeline [43], based on 44 haplotype genomes derived from 22 accessions (excluding the 'Donghong' and 'wild' accessions due to their mosaic genome assemblies) [23, 41]. The graph pangenome constructed is 2.24 Gb in size, 3.6 times as large as that of the MDHAPA genome (Additional file 1: Table S6) [31]. A total of 31,790,044 SNPs, 13,512,079 InDels (< 50 bp) and 623,478 SVs (≥ 50 bp) were identified (Fig. 1c). The size of most SVs is smaller than 2 Kb (Fig. 1d). These SVs constitute a total of 1.5 Gb of sequences, with an average of 34.1 Mb per genome. In addition, we found that approximately 48.05% of SVs overlap with TEs' locations (Fig. 1e), suggesting that TEs might be a significant driving force for the formation of SVs. Out of these SVs, 269,784 SVs are highly differentiated between *A. chinensis* var. *chinensis* and *A. eriantha* (Fig. 1f), which likely represent important genomic determinants underlying the interspecific differentiation between the two species.

Construction of pan-NLRome

The nucleotide-binding domain and leucine-rich repeat (NLR) proteins have been shown to play a crucial role in plant immunity by specifically recognizing pathogen effectors [44]. To explore the relationship between NLR genes and immunity

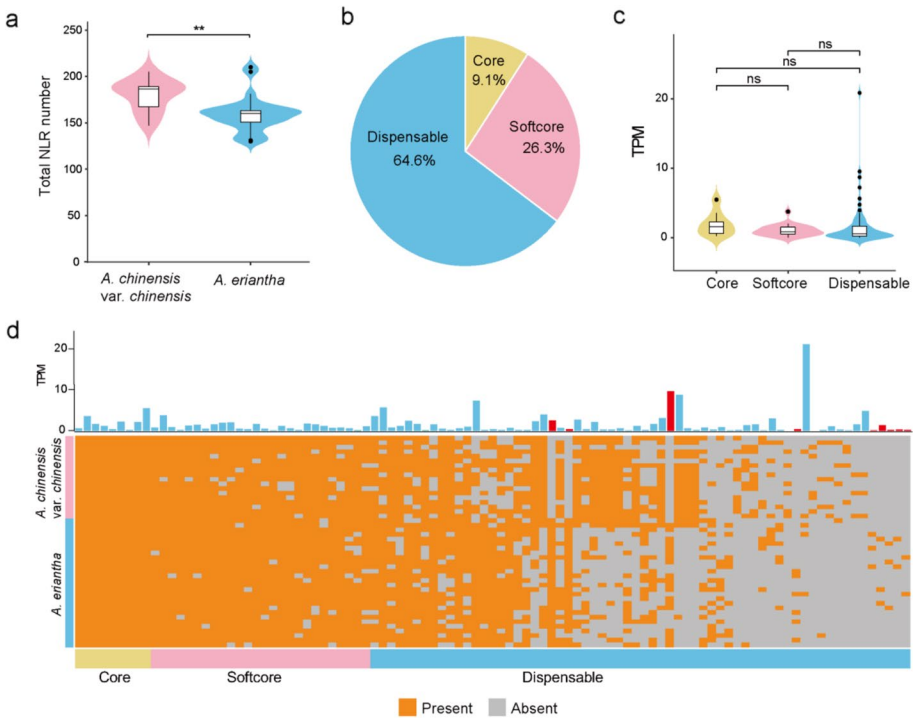

Fig. 2 Characterization of the pan-NLRome in pangenome across *A. eriantha* and *A. chinensis* var. *chinensis*. **a** Comparison of the total number of NLR genes between *A. eriantha* and *A. chinensis* var. *chinensis*. **b** The percentage of core, softcore, and dispensable NLR orthogroups in the pan-NLRome. **c** Comparison of the expression values among core, softcore, and dispensable NLR gene orthogroups. **d** Presence and absence of the NLR gene orthogroups in *A. chinensis* var. *chinensis* and/or *A. eriantha*. The bars in the upper panel represent the expression levels of different non-redundant NLR gene orthogroups. Red bars indicate NLR gene orthogroups exclusive to *A. eriantha* and absent in all accessions of *A. chinensis* var. *chinensis*

difference, we identified 7,692 NLR genes across the 46 assemblies, with significant variation in the number of NLR genes among different accessions, ranging from 130 to 210 (Fig. 2a, Additional file 1: Tables S7 and S8). On the basis of domain and motif structures [45], the identified NLRs were classified into six categories: TN (Toll/interleukin-1 receptor (TIR) and NB-ARC), CN (coiled-coil (CC) and NB-ARC), NL (NB-ARC and leucine-rich repeat (LRR)), CNL (CC, NB-ARC, and LRR), NB (NB-ARC), and TNL (TIR, NB-ARC, and LRR) (Additional file 1: Table S7). We found that the number of NLR genes in *A. chinensis* var. *chinensis* was slightly larger than in *A. eriantha* (Fig. 2a). In fact, larger number of N, CN, and CNL-type NLR genes in *A. chinensis* var. *chinensis* was present than in *A. eriantha*, whereas significantly smaller number of TNL-type NLR genes was annotated in *A. chinensis* var. *chinensis* compared to *A. eriantha* (Fig. S7a). If the NB group is not included, we found that the total NLR counts in *A. eriantha* and *A. chinensis* var. *chinensis* showed no significant difference, but that *A. eriantha* exhibited significantly greater variation than *A. chinensis* var. *chinensis* (Additional file 2: Fig. S7b).

Next, we used the OrthoFinder software [46] to cluster the 7,692 NLR genes, and 99 non-redundant NLR orthogroups were identified (Additional file 1: Table S9). In the constructed pan-NLRome, core NLR orthogroups account for 9.1%, softcore NLR orthogroups for 26.3%, and dispensable NLR orthogroups for 64.6% (Fig. 2b). Then

we selected leaves from 16 accessions (eight from either *A. chinensis* var. *chinensis* or *A. eriantha*, see materials and methods) for transcriptome sequencing. The expression levels of core, softcore, and dispensable NLR orthogroups showed no significant differences (Fig. 2c), but several NLR orthogroups belonging to the dispensable group exhibited exceptionally high expression (Fig. 2d). Most NLRs are shared by both *A. eriantha* and *A. chinensis* var. *chinensis* accessions, and a total of nine NLR orthogroups that are present only in *A. eriantha* and absent in all accessions of *A. chinensis* var. *chinensis* (Fig. 2d). In addition, genomic comparison analysis revealed that 448 out of 269,784 SVs highly differentiated between *A. chinensis* var. *chinensis* and *A. eriantha* are associated with 69 NLR genes (Additional file 1: Table S10).

Diverged *Psa* resistance between *A. chinensis* var. *chinensis* and *A. eriantha* and species-specific SV associated with allele-specific expression of *NLR25-1* in the interspecific F1 hybrid

It has been known that the vast majority of *A. chinensis* var. *chinensis* varieties are susceptible to *Psa*, while *A. eriantha* accessions are resistant [47, 48]. To explore the genetic basis of the differential immune abilities against *Psa*, an interspecific hybridization was made using an *A. eriantha* 'White' [33] as female pollinated by a male *A. chinensis* var. *chinensis* 'H0809' [35]. A female F1 individual (named 'HH-1') and its parents were employed to evaluate the phenotypic performance in response to *Psa*-inoculation. Their leaf discs or shoots at uniformly developmental stages were inoculated with *Psa* race 'ScBcH' isolated previously from the infected field plants of *A. chinensis* var. *chinensis* cv. 'Hongyang' [49]. Disease symptoms appeared in leaf discs at three days post-inoculation (dpi) or shoots at 21 dpi were observed (Fig. 3a) and evaluated by measuring lesion size (Fig. 3b) or length (Fig. 3c) of the infected tissues. This analysis showed that 'H0809' tissues were extensively infected, while 'White' tissues remained almost innate immune, and their hybrid was in an intermediate state. In addition, by using leaf discs inoculation with *Psa* race 'ScBcH', we also evaluated the phenotypic performance of *Psa* resistance/susceptibility for all other accessions employed for pangenome construction at 3 dpi. As

(See figure on next page.)

Fig. 3 Structural variation (SV) in promoter led to allele-specific expression (ASE) at *NLR25-1* locus in F1 hybrid between *A. eriantha* 'White' and *A. chinensis* var. *chinensis* 'H0809'. **a**-**c** The phenotypic symptom (**a**) and quantification through measuring lesion area in leaf discs (**b**) or lesion length in shoots (**c**) of 'White', 'H0809', and 'HH-1' at 3 or 21 dpi with *Psa*. Each data point represents the mean (± SD) of three replications, and asterisk refers to significant difference (Student's t-test; **: $p < 0.01$; ***: $p < 0.001$). **d** The gene counts of allele-specific expression genes (ASEGs) of 'HH-1' upon *Psa*-inoculation at different hpi. ASEG1s or ASEG2s indicate gene counts that are preferentially expressed in HH-1.hap1 or HH-1.hap2. The blue or pink bar represents non-pathogen-resistant or pathogen-resistant genes in ASEGs, respectively. **e** Expression profiles of 87 allelic gene pairs with ASE pattern in individual haplotypes of 'HH-1' upon *Psa*-inoculation. **f** Expression patterns of *NLR25-1* in 'HH-1' and its parents (**g**) PCR validation of allelic structure variation. **h** Diagram of structural variation at *NLR25-1* locus of 'HH-1'. *AeNLR25-1* is derived from HH-1.hap1 (inherited from *A. eriantha* 'White'), while *AcNLR25-1* is derived from HH-1.hap2 (inherited from *A. chinensis* var. *chinensis* 'H0809'). Blue indicates syntenic regions upstream the start codon (ATG), while yellow represents the structural variation (SV) derived from individual haplotypes. The top or bottom black solid lines represent upstream promoter sequences that are designated as *ProAeNLR25-1* or *ProAcNLR25-1*, respectively. The red, blue, green, and dark green vertical lines on the black horizontal lines corresponding to the conserved binding motifs recognized by WRKY (binding to TGACC/T), MYB (binding to TAACTG/CAACTG), bHLH (binding to CAGGTG/CACGYG), and AP2 (binding to GCCGCC/ACCGC/CCGAC) transcription factors, respectively. All the experiments and measurements have been independently performed with three biological replicates and three technical replicates

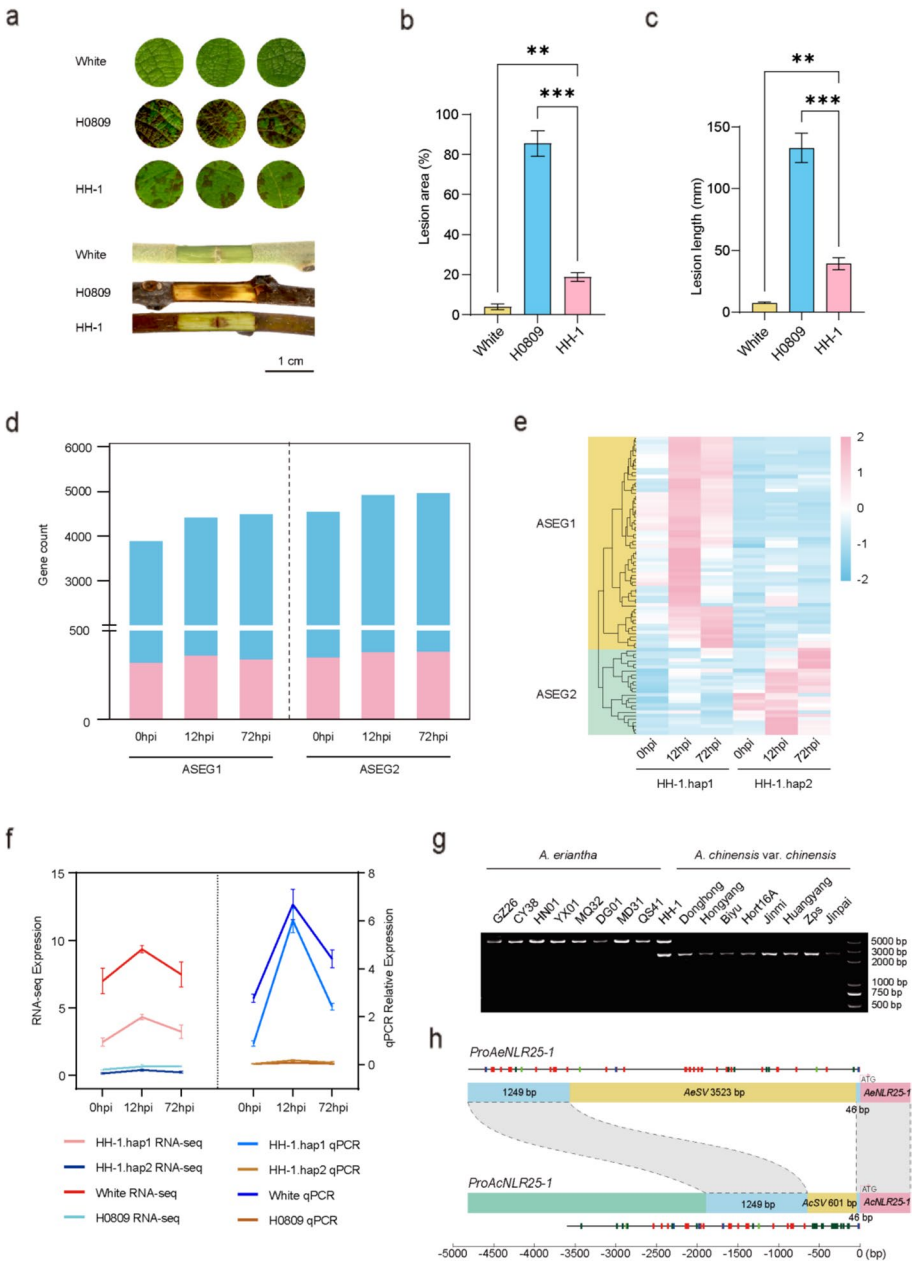

Fig. 3 (See legend on previous page.)

a result, all tested *A. eriantha* accessions were highly resistant to the pathogen, while the *A. chinensis* var. *chinensis* accessions displayed different magnitudes of susceptibility (Additional file 2: Fig. S8).

Incorporating beneficial alleles or haplotypes from wild relatives into breeding program is a crucial step for crop improvement [50]. To this end, identification of allele-specific expression (ASE, referring to the preferential expression of a parental allele in the hybrid due to variations in regulatory sequences from the parental genomes) is essential [51, 52]. To explore ASE in the *A. eriantha* × *A. chinensis* var. *chinensis* F1 hybrid, we performed transcriptome sequencing on the leaves of 'HH-1' at 0-, 12-, and

72-h post-inoculation (hpi) with *Psa* race 'ScBcH' (Additional file 1: Tables S11 and S12). As compared to the control (0 hpi), 24,361 (11,932 for HH-1.hap1 and 12,429 for HH-1.hap2) and 28,694 (14,047 for HH-1.hap1 and 14,647 for HH-1.hap2) differentially expressed genes (DEGs) were identified in samples at 12 and 72 hpi (Additional file 1: Tables S13 and S14), respectively. Using JCVI [53], 27,553 pairs of allelic genes were identified between HH-1.hap1and HH-1.hap2 (Additional file 1: Tables S15 and S16). A total of 8,649, 9,546 and 9,717 pairs of significant ASE genes (ASEGs) were filtered in samples at 0, 12 and 72 hpi, respectively (Additional file 1: Tables S17, S18 and S19). Subsequently, 3,139 or 2,985 pathogen-resistant genes including salicylic acid pathway related genes and immune receptor-like genes (Additional file 1: Table S20) were annotated in HH-1.hap1 or HH-1.hap2 utilizing DRAGO and HMMER respectively [54, 55]. Among ASEGs, 3,999, 4,520 and 4,589 ASEGs were biased expressors of HH-1.hap1 derived from *A. eriantha* (designated as ASEG1) at 0, 12 and 72 hpi, respectively, including 321, 364 and 340 pathogen-resistant genes (Fig. 3d, Additional file 1: Table S21). And 62 allelic pathogen resistant gene pairs with a significant up-regulation after *Psa*-inoculation and a continuous expression bias to HH-1.hap1 were identified (Fig. 3e, Additional file 1: Table S22). Meanwhile, 4,650, 5,026 and 5,128 gene pairs were biased expressors of HH-1.hap2 derived from *A. chinensis* var. *chinensis* (designated as ASEG2) at 0, 12 and 72 hpi, respectively, including 348, 377 and 377 pathogen-resistant genes (Fig. 3d, Additional file 1: Table S23). And 25 allelic pathogen-resistant gene pairs with a significant up-regulation after the *Psa* inoculation and a continuous expression bias to HH-1.hap2 were identified (Fig. 3e, Additional file 1: Table S24). Similar analyses were conducted on the parental tissues to explore the differently expressed orthologous genes at 0, 12 and 72 hpi with *Psa*-inoculation (Additional file 1: Tables S25, S26 and S27). Consistently, transcriptomic profile comparison between the parental lines showed that expression of over 70% ASEG1 orthologs in 'White' was significantly higher than that of 'H0809' and that expression of over 70% ASEG2 orthologs in 'H0809' was significantly higher than that of 'White' (Additional file 2: Fig. S9).

Out of ASEG1s with a significant up-regulation after the *Psa*-inoculation, a CNL member from the NLR gene family, designated as *NLR25-1* located on chromosome 25 with a pair of allelic genes (*AeNLR25-1/Aisf1h1a25g408720* and *AcNLR25-1/Aisf1h1b25g410370*), was resident in an interspecies-specific SV (Table S10) and showed extremely biased expression toward HH-1.hap1 in hybrid 'HH-1' (Fig. 3f, Additional file 1: Table S16). To validate the allele-specific expression pattern at the *NLR25-1* locus, qRT-PCR analysis was performed on the *Psa*-inoculated leaves of 'HH-1', 'H0809' and 'White' at 0, 12 or 72 hpi. Consistently, abundant *AeNLR25-1* transcripts were detected at 0 hpi and its expression was strongly reinforced at 12 and 72 hpi, while *AcNLR25-1* transcripts were hardly detected either before or after *Psa*-inoculation (Fig. 3f). However, both *AeNLR25-1* and *AcNLR25-1* encode a protein with identical size of 1040 amino acids, and thirty-four SNPs are present across the allelic coding sequences, leading to 20 amino acid substitutions (Additional file 2: Fig. S10).

To dissect the mechanism underlying the allele-specific expression at the *NLR25-1* locus, our subsequent investigation focused on the structure variations in their promoter regions. Consistent with the SVs detected by pangenome analysis (Additional file 1: Table S10), sequence comparison between the allelic regions revealed a conserved

interspecific SV resident in the upstream regulatory region at the *NLR25-1* locus, i.e. unique fragments ranging from 3,409 to 3,720 bp were consistently identified at 46 bp upstream of the *AeNLR25-1* gene's translation start site across individual *A. eriantha* accessions, whereas this region was replaced by distinct sequences ranging from 549 to 622 bp in individual *A. chinensis* var. *chinensis* accessions, creating a significant interspecies divergence in the *NLR25-1* promoter architecture (Additional file 2: Fig. S11a). In accordance with pangenome analysis, this difference is confirmed in multiple *A. chinensis* var. *chinensis* and *A. eriantha* accessions through PCR amplification and agarose gel electrophoresis, in which all *A. eriantha* accessions showed a larger amplified fragment than from *A. chinensis* var. *chinensis* (Fig. 3g). More specifically, a unique 3,523 bp sequence (namely *AeSV*) was identified at 46 bp upstream of the *AeNLR25-1* translation start site in 'White' and HH-1.hap1 from 'HH-1', whereas this region was replaced by a distinct 601 bp fragment (namely *AcSV*) in 'H0809' and HH-1.hap2 from 'HH-1' (Fig. 3h). Further investigation revealed that the sequence divergence is primarily due to the insertion of six additional transposable elements (TEs), including one LTR and five terminal inverted repeats, in the promoter region of *AeNLR25-1* compared to that of *AcNLR25-1* with a single terminal inverted repeat (Additional file 2: Fig. S11b).

NLR25-1 functions in immune response against *Psa*

To unravel the biological function of NLR25-1, CRISPR-cas9-mediated mutants of 'HH-1' were constructed by *Agrobacterium*-mediated genetic transformation, and three independent *NLR25-1*-knockout lines (*nlr25-1–1*/*nlr25-1–2/nlr25-1–3*) were obtained (Additional file 2: Fig. S12). Then leaf discs, tissue culture seedlings with uniform developmental stage, or 5-leaf-age potted seedlings derived from wild-type 'HH-1' (WT) or three mutants were inoculated with *Psa* race 'ScBcH'. Increased lesion area was observed on leaf discs of mutant lines compared to WT at 3 dpi (Additional file 2: Fig. S13a and S13b). The average *in planta* bacterial growth of *Psa* bacteria in tissue culture seedlings was higher in mutants compared to WT at 7 dpi (Additional file 2: Fig. S13c and S13d). The potted seedlings of mutant lines displayed a severe wilting phenotype with

(See figure on next page.)

Fig. 4 Functional characterization of *NLR25-1* using knockout mutants or overexpression transgenics in responses to *Pseudomonas syringae pv. actinidiae* (*Psa*) inoculation. **a** Phenotypic comparison of 'HH-1' and its CRISPR-cas9-mediated *NLR25-1*-knockout plants in potted seedlings at 7 dpi (days post inoculation) with *Psa* ('ScBcH') (lower panel). The upper panel represents control without *Psa*-inoculation. **b** The *in planta* bacterial growth measured for the *Psa*-inoculated seedlings in (**a**). **c** Staining of leaves from *Psa*-inoculated seedlings at 1 or 7 dpi with trypan blue to assess cell death, or NBT and DAB to assess superoxide production, respectively. The depth of color represents the quantity or content. **d**-**f** Physiological changes measured for the *Psa*-inoculated seedlings at 1 or 7 dpi, including superoxide (O_2^-) content (**d**), hydrogen peroxide (H_2O_2) content (**e**), malondialdehyde (MDA) content (**f**). **g** Phenotypic comparison of 'H0809' and *NLR25-1*-OE plants in potted seedlings at 7 dpi with *Psa* ('ScBcH') (lower panel). The upper panel represents control without *Psa*-inoculation. **h** The *in planta* bacterial growth measured for the *Psa*-inoculated seedlings in (**g**). **i** Staining of leaves from *Psa*-inoculated seedlings at 1 or 7 dpi with Trypan blue to assess cell death, or NBT and DAB to assess superoxide production, respectively. The depth of color represents the quantity or content. **j**-**l** Physiological changes measured for the *Psa*-inoculated seedlings at 1 or 7 dpi, including superoxide (O_2^-) content (**j**), hydrogen peroxide (H_2O_2) content (**k**), malondialdehyde (MDA) content (**l**). Each data point represents the mean (± SD) of three replications. Different letters on error bar referred to significant difference based on one-way ANOVA and Tukey's test ($p < 0.05$). All the experiments and measurements have been independently performed with three biological replicates and three technical replicates

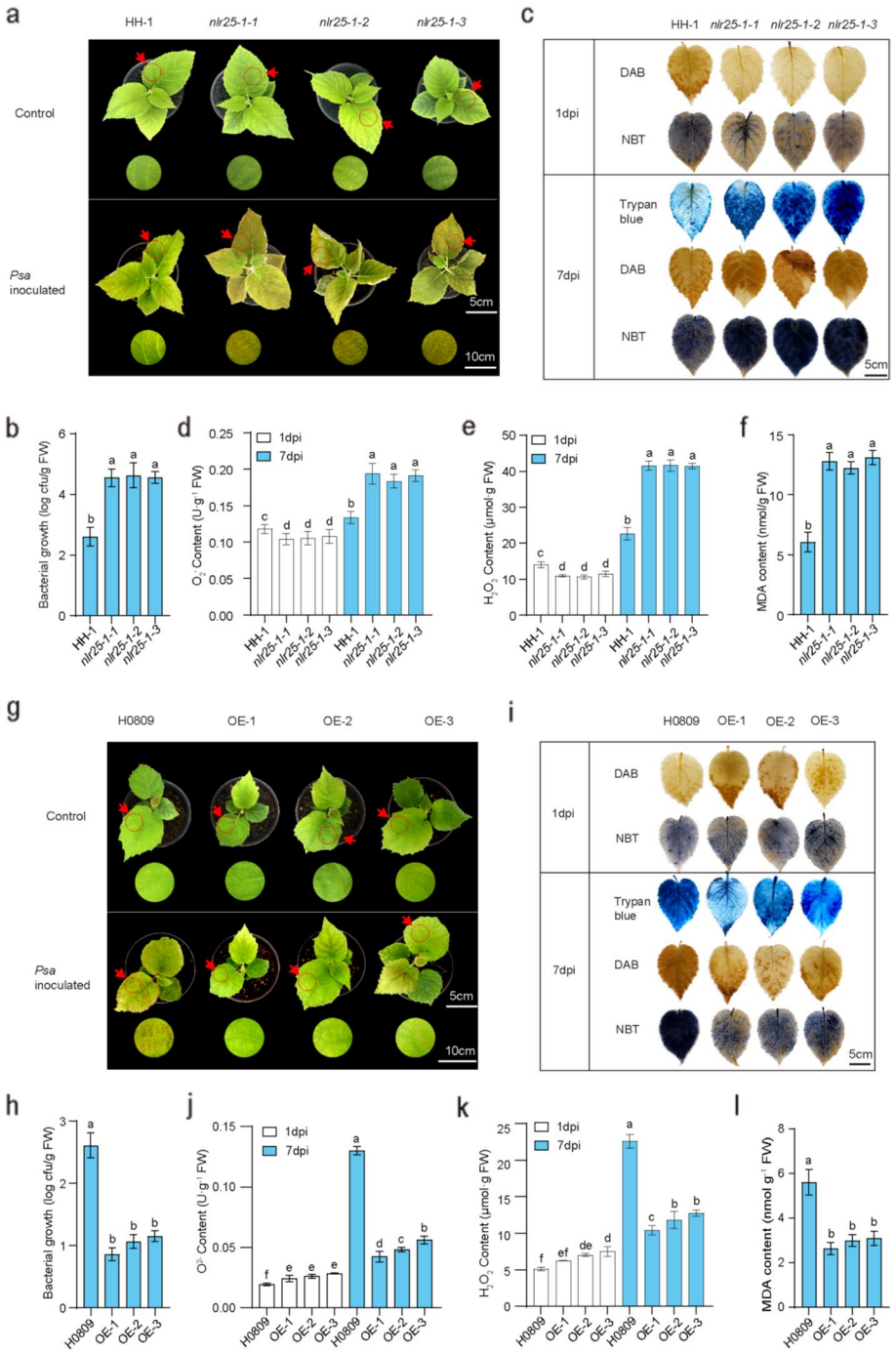

Fig. 4 (See legend on previous page.)

almost entire leaf areas of expanded lesion compared to WT that exhibited mild wilting symptoms with restricted area of lesion expansion (Fig. 4a), and the average *in planta* bacterial growth in WT was less than in the mutants at 7 dpi (Fig. 4b). NBT staining, DAB staining [56, 57], as well as O_2^- and H_2O_2 quantification were conducted on leaf samples of the potted seedlings at 1 dpi to detect the immunity performance, and the results showed that the immunity responses of mutants were significantly weaker than

WT (Fig. 4c-e). In addition, to further assess the disease symptom, leaf samples of the potted seedlings at 7 dpi were collected and subjected to physiological index measurement. Trypan blue staining, NBT staining and DAB staining [56, 57] was conducted, and the results showed that dead cells' number and hydrogen peroxide content were significantly increased in mutant leaves (Fig. 4c). Consistently, the average contents of O_2^-, H_2O_2 or MDA significantly increased in mutant lines compared to WT (Fig. 4d-f), and the activities of defense enzymes including catalase (CAT), peroxidase (POD) or superoxide dismutase (SOD) in WT were significantly higher than mutants (Additional file 2: Fig. S13e-g). These results indicated that the *nlr25-1* mutants displayed more severe cellular damage (necrosis) resulting from *Psa*-mediated abiotic stress than WT. Meanwhile, the expression of defense-related genes from salicylic acid (SA) pathway including homologs of *PR1* (*Aisf1h1a19g304391/Aisf1h1b19g310001*, Additional file 2: Fig. S13h), *TGA9* (*Aisf1h1a12g195220/Aisf1h1b12g198870*, Additional file 2: Fig. S13i), *PR9* (*Aisf1h1a28g469250/Aisf1h1b28g469530*, Additional file 2: Fig. S13j) and *NPR1* (*Aisf1h1a09g139020/Aisf1h1b09g139210*, Additional file 2: Fig. S13k) was detected at 0 or 72 hpi through qRT-PCR analysis, and the result showed that all the tested genes were significantly induced and the reinforced expression magnitude in WT was much higher than in mutant plants. Taken together, these results suggested that *NLR25-1* plays a vital role in resistance to *Psa*.

To further assess the immune activity of NLR25-1, three independent transgenic lines constitutively overexpressing *AeNLR25-1* (OE1/OE2/OE3) were obtained through *Agrobacterium*-mediated genetic transformation using 'H0809' as transgene host (Additional file 2: Fig. S14a) [58]. Upon *Psa* inoculation, leaf discs at 3 dpi showed reduced lesion area in three OE lines compared to WT (Additional file 2: Fig. S14b and S14c). Similarly, *in planta* bacterial growth of *Psa* bacteria for tissue culture seedlings or potted seedlings in OE lines at 7 dpi were significantly lower than that of WT (Fig. 4g, h, Additional file 2: Fig. S13d and S13e). NBT staining, DAB staining, as well as O_2^- and H_2O_2 quantification were conducted on leaf samples of the potted seedlings at 1 dpi to measure the immunity response, and the results showed that the immunity responses of OE lines were stronger than WT (Fig. 4i-k). Disease severity was further accessed through physiological index measurement for potted seedlings at 7 dpi. Dead cells and hydrogen peroxide were reduced in OE leaves compared to WT (Fig. 4i). The contents of O_2^-, H_2O_2 and MDA in WT were significantly higher than that in OE lines (Fig. 4j-l), while the activities of CAT, POD and SOD in WT were significantly lower than that in OE lines (Additional file 2: Fig. S14f-h). Consistently, the defense genes' transcription levels were significantly higher in OE plants after *Psa* inoculation (Additional file 2: Fig. S14i-l). These results further suggest a crucial role conferred by *AeNLR25-1* in resistance to *Psa*.

To provide more evidence supporting that the SV-associated *AeNLR25-1* confers the resistance to *Psa*, a backcross population (BCF1) was constructed using 'HH-1' as female pollinated by the male parent 'H0809'. Leaf tissues from 120 BCF1 offspring were challenged by the *Psa* infection and subjected to phenotypic assessments (Additional file 2: Fig. S15a). To determine the genotype of the tested BCF1 plants, a pair of distinct PCR primers was designed to distinguish the allelic difference at the *NLR25-1* locus (Additional file 2: Fig. S15b). Genotyping was then conducted on the 120 plants, resulting in 44 heterozygotes (*AeNLR25-1/AcNLR25-1*) and 76 homozygotes (*AcNLR25-1/AcNLR25-1*)

(Additional file 2: Fig. S15c). The phenotypic analysis showed that the average lesion area of homozygous or heterozygous plants was 12.10% or 7.32% (Additional file 2: Fig. S15d), respectively, indicating that the heterozygous plants were more resistant to *Psa*-inoculation and further supporting the immune function by *AeNLR25-1* preferential expression in response to *Psa*-inoculation.

Both AeWRKY75 and AcWRKY75 specifically bind to *AeSV* and activate *AeNLR25-1* transcription and immune responses

To study how the identified SV affects the activity of promoter of *AeNLR25-1* (*ProAeNLR25-1*) or *AcNLR25-1* (*ProAcNLR25-1*), a GUS staining assay was conducted in 'HH-1' leaves by *Agrobacterium*-mediated infiltration, and the result indicated that the activity of *ProAeNLR25-1* was much stronger than *ProAcNLR25-1* (Fig. 5a). These results are consistent with the transcriptome profile dataset and qRT-PCR analysis (Fig. 3e and f).

To explore the underlying mechanism of differential activities between the allelic promoters (*ProAeNLR25-1*/*ProAcNLR25-1*), a total of 2,632 or 2,569 transcription factors were predicted in HH-1.hap1 or HH-1.hap2, respectively (Additional file 1: Table S28), and the transcriptome profiles with 5,201 annotated transcription factors in 'HH-1' genome were inspected. In 'HH-1' samples at 12 hpi with *Psa* when *AeNLR25-1* expressed at the most abundant level (Fig. 3f), we found a total of 841 transcription factors were significantly upregulated (Additional file 1: Table S29) and enriched in WRKY transcription factor family (Additional file 2: Fig. S16). The twelve WRKY members with top Pearson correlation coefficient in expression with *NLR25-1* were screened as candidates (Additional file 1: Table S30) and validated utilizing Dual-LUC assay, and one of them, namely AeWRKY75 (*Aisf1h1a29g472580*), was eventually verified to be able to activate the promoter of *AeNLR25-1* (Additional file 2: Fig. S17).

To demonstrate the function of WRKY75 homologs in *Psa* resistance, the allelic gene of *AeWRKY75*, namely *AcWRKY75* (*Aisf1h1b29g473170*) was also identified and

(See figure on next page.)

Fig. 5 AeWRKY75/AcWRKY75 specifically binds to and activates the promoter of *AeNLR25-1* and confers resistance to *Pseudomonas syringae pv. actinidiae* (*Psa*). **a** GUS staining assay validating promoter activity at the *NLR25-1* locus by detecting differential GUS activities in 'HH-1' leaves transiently expressing different constructs. The expression level in *ProAeNLR25-1*::GUS was designated as "1". **b** Dual-LUC assay validating activity of AeWRKY75/AcWRKY75 by activating *AeProNLR25-1*/*AcProNLR25-1* in *Nicotiana benthamiana* leaf (**: $p < 0.01$; ns: not significant). **c** Y1H assay showing binding ability of AeWRKY75/AcWRKY75 to the promoters (*ProAeNLR25-1*/*ProAcNLR25-1*). The putative binding sites of WRKY family transcription factors on truncated segments (*AeP2.1*, *AeP2.2* and *AeP2.3*) are marked by red vertical lines. The blue colony in lower panel shows the ability of binding. **d** The electrophoretic mobility shift assay (EMSA) validating binding ability of AeWRKY75 and AcWRKY75 to *AeSV*/*ProAeNLR25-1*. **e** The expression alteration of *AeNLR25-1*/*AcNLR25-1* in 'HH-1' or *WRKY75*-knockout plants with *Psa*-inoculation. **f** Phenotypic comparison of 'HH-1' and *WRKY75*-knockout plants (potted seedlings) at 7 days post inoculation (dpi) (lower panel). The upper panel represents control without *Psa*-inoculation. **g** The *in planta* bacterial growth measured for the *Psa*-inoculated plants in (**f**). **h** Staining of leaves from *Psa*-inoculated seedlings at 1 or 7 dpi with trypan blue to assess cell death, or NBT and DAB to assess superoxide production, respectively. The depth of color represents the quantity or content. **i**-**k** Physiological changes in 'HH-1' and mutant plants at 1 or 7 dpi, including superoxide (O_2^-) content (**i**), hydrogen peroxide (H_2O_2) content (**j**), malondialdehyde (MDA) content (**k**). Each data point represents the mean (± SD) of three replications. Different letters on error bar referred to significant difference based on one-way ANOVA and Tukey's test ($p < 0.05$). All the experiments and measurements have been independently performed with three biological replicates and three technical replicates

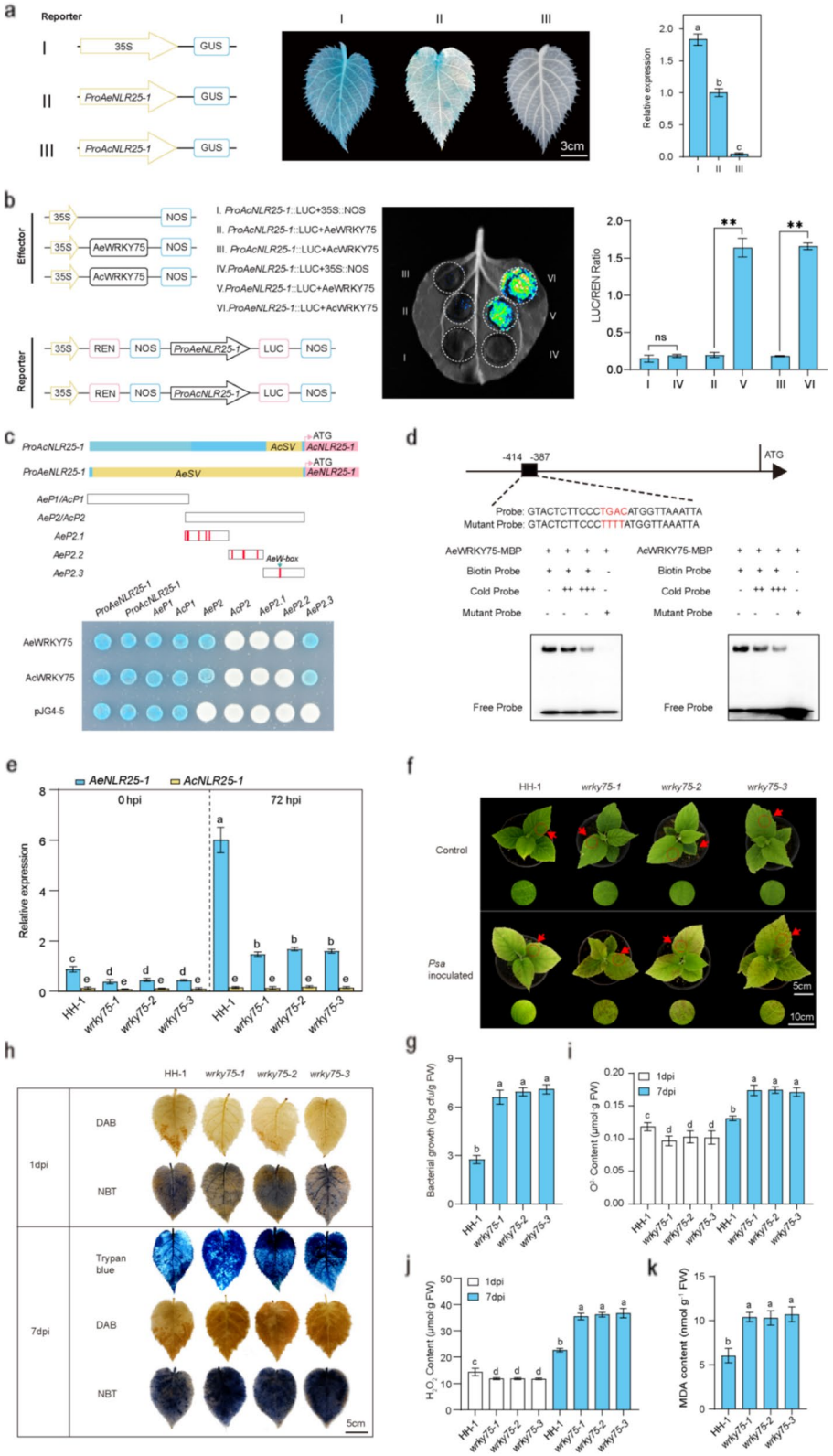

Fig. 5 (See legend on previous page.)

isolated (Additional file 2: Fig. S18a and S18b). Moreover, combined with the transcriptome profile analysis and qRT-PCR analysis, without *Psa*-inoculation, both *AeWRKY75* and *AcWRKY75* showed similarly low level of expression. By contrast, both *AeWRKY75* and *AcWRKY75* transcripts were abundantly induced upon *Psa*-inoculation (Additional file 2: Fig. S18c). Next, dual-luciferase reporter assay suggested either AeWRKY75 or AcWRKY75 is able to activate the *AeNLR25-1* promoter containing *AeSV* but neither of them can activate the *AcNLR25-1* promoter containing *AcSV* (Fig. 5b). Yeast-one-hybrid (Y1H) assay indicated that the binding site was present in the *AeP2.3* fragment derived from *AeSV* (Fig. 5c). The electrophoresis mobility shift assay (EMSA) showed a *W-box* motif (TGAC/GTCA) in the *AeP2.3* fragment is the precise binding site (namely *AeW-box*) (Fig. 5d). Although autoactivation was observed on *AeP1* and *AcP1*, further EMSA assays ruled out the possibility of additional functional binding site(s) available for WRKY75 binding (Additional file 2: Fig. S19). These results suggest that both AeWRKY75 and AcWRKY75 specifically binds to *AeW-box* in *AeSV*/*ProAeNLR25-1* to activate *AeNLR25-1* expression, while the *AcNLR25-1* cannot be activated due to the *ProAcNLR25-1* lacking the corresponding *cis*-element.

To validate the possible function of *WRKY75* in *Psa* resistance, CRISPR-cas9-mediated mutants of 'HH-1' (*AeWRKY75*/*AcWRKY75*) were constructed by *Agrobacterium*-mediated genetic transformation [58], and three independent knockout lines (*wrky75-1/2/3*) were obtained (Additional file 2: Fig. S20). Then, we evaluated the expression level of *NLR25-1* in *wrky75* mutants. qRT-PCR analysis showed that the expression level of *AeNLR25-1* was significantly lower in mutant plants than WT without *Psa*-inoculation (Fig. 5e). Moreover, the expression of *AeNLR25-1* was significantly induced upon *Psa*-inoculation in both WT and mutants, but the increased level in mutant plants was much lower than WT. By contrast, the expression level of *AcNLR25-1* in *wrky75-1*/*2*/*3* mutants and WT plants remained low level or unchanged before or after *Psa*-inoculation. These together suggest that AeWRKY75/AcWRKY75 plays a critical role in *Psa* resistance probably via binding to *AeW-box* and activating *AeNLR25-1* expression.

Meanwhile, phenotypic evaluation showed larger lesion area was observed on leaf discs of mutants than WT upon *Psa*-inoculation at 3 dpi (Additional file 2: Fig. S21a and S21b). The *in planta* bacterial growth in the treated seedlings is also higher in mutants than WT at 7 dpi (Fig. 5f, g, Additional file 2: Fig. S21c and S21d). NBT staining, DAB staining, as well as O_2^- and H_2O_2 quantification were conducted on leaf samples of the potted seedlings at 1 dpi to detect the immunity response, and the results showed that the immunity responses of mutants were weaker than WT (Fig. 5h-j). Disease severity was further accessed through physiological index measurement for potted seedlings at 7 dpi. The dead cells and hydrogen peroxide were significantly increased in mutant leaves compared to WT (Fig. 5h). We also measured the contents of O_2^-, H_2O_2 and MDA or activities of POD, SOD and CAT. The results indicated that the average contents of O_2^-, H_2O_2 and MDA in the mutants were significantly increased compared to WT (Fig. 5i-k), while the average activities of CAT, POD, and SOD in mutants were significantly decreased compared to WT (Additional file 2: Fig. S21e-g). In addition, the expression of defense-related genes was also tested. All the tested genes were expressed at a low level in both WT and the three mutant lines at 0 hpi, while their expression was substantially

induced at 72 hpi, but the induction was substantially inhibited by *AeWRKY75/AcWRKY75*-defective mutations compared to WT (Additional file 2: Fig. S21h-k).

Discussion

Despite the protracted process that canonical crops underwent from wild forms to domesticated forms, accumulating evidence suggests that the rapid reduction in diversity of modern cultivated varieties is a result of modern breeding and cultivation practices [59–61]. This is exactly the case for kiwifruit whose domestication only began in the early twentieth century and no more than 100 years of successive rounds of artificial selection has generated numerous distinctive cultivars, forming a huge kiwifruit industry worldwide [1]. However, after decades of intensive cultivation, these domesticates are highly susceptible to canker pathogens of *Psa*. In the present study, a reference-unbiased pangenome construction allowed identification of 623,487 SVs. Our genomic data indicates that TE-induced genome size expansion in the wild *A. eriantha* accessions may also contribute to the increased SVs number and genomic diversity. We combined pangenome data with transcriptome analysis of interspecific samples between the cultivated individual and the wild relative to discover the association between many allele-specific expression (ASE) gene loci and SVs. What is particularly interesting is the drastic changes in expression patterns caused by the highly differentiated regulatory regions of disease resistant genes.

To confront the distinct defense abilities against *Psa* between the cultivated *A. chinensis* var. *chinensis* and wild *A. eriantha*, we characterized the ASE profile of all annotated NLR genes that can be paired one on one. A total of 76 pairs of allelic NLR genes were matched, and 52 pairs were validated as ASE genes for at least one timepoint among 0, 12 or 72 hpi in our transcriptomic data (Additional file 1: Table S31). Considering the abundance of structure variations between the two species, the upstream 5,000 bp of each pair of allelic NLR genes were retrieved and aligned. Thirty-one pairs of allelic NLR genes showed low identity in upstream region, while 20 or 17 pairs showed small or large segment alignment, and 8 pairs showed a highly matched upstream sequence similarity. The proportion of ASE genes increased with the difference of upstream sequences enlarged (Additional file 2: Fig. S22). Further investigation showed that the large variations were owing to multiple TE insertions (Additional file 1: Table S31).

We conducted a detailed genetic and molecular analysis of this change at the *NLR25-1* locus that was initially identified through transcriptome profiling and ASE analysis (Fig. 3e-f; Additional file 1: Table S16). Our analysis indicated that it is a typical ASE gene and its regulatory region was associated with a TE-induced SV whose length and structure are highly conserved within parental accessions. We provide evidence showing the wild allele *AeNLR25-1* is predominantly expressed in the hybrid and confers resistance to *Psa*. We discovered that the preferential expression of *AeNLR25-1* in response to *Psa*-inoculation is owing to the presence of a unique *A. eriantha*-specific *AeW-box* (TGAC/GTCA) that could be bound and activated specifically by WRKY transcription factors AeWRKY75 or AcWRKY75.

To inspect the specificity of the conserved structure variation at the *NLR25-1* locus, a total of 75 sequenced accessions of various *Actinidia* species were examined including all 46 accessions in the interspecific pan-genomics in this study, as well as 29 accessions

retrieved from Kiwifruit Pan-Genome Database (KPGD) [62] consisting of two *A. eriantha*, one *A. chinensis* var. *chinensis*, nine *A. arguta*, six *A. chinensis* var. *deliciosa*, one *A. hemsleyana*, one *A. latifolia*, one *A. longicarpa*, one *A. macrosperma*, two *A. polygama*, one *A. reticulata*, two *A. rufa* and two *A. zhejiangensis* (Additional file 1: Table S32). We found only a single allele at the *NLR25-1* locus was missing in *A. chinensis* var. *deliciosa* haplotype 'MeiweiW2.hap5' and *A. chinensis* var. *chinensis* haplotype 'Jinpai.hap2', suggesting a highly conserved existence of the *NLR25-1* locus. Importantly, the *AeW-box cis*-element is exclusively present in all tested *A. eriantha* assemblies and an *A. zhejiangensis* haplotype 'ZhejiangW1.hap1', which has been postulated as a hybrid derived from a cross between *A. eriantha* and *A. hemsleyana* [63], demonstrating the specificity of *AeW-box cis*-element uniquely existing in *A. eriantha* (Additional file 2: Fig. S23, Additional file 1: Table S32). Interestingly, a mutation on the translation start site of the *NLR25-1* locus occurs in five accessions including *A. arguta* accessions 'Longcheng2V1.hap3' and 'RuanzaoW2.hap1', *A. chinensis* var. *deliciosa* accession 'MeiweiW2.hap1', *A. hemsleyana* accession 'ChangyeW1' and *A. zhejiangensis* 'ZhejiangW1.hap2', probably resulting in a delayed translation initiation at 88 bp, causing a skip of 29 amino acid in the translation products. Furthermore, all *A. arguta* accession consists of one or two duplicated segments sized 81 bp at the *NLR25-1* locus (Additional file 2: Fig. S24 and S25, Additional file 1: Table S32). However, these structural variations do not alter regions coding for either CC, NB-ARC or LRR domains at the *NLR25-1* locus.

It is worth noting that although *NLR25-1* gene was not detected in haplotype 2 of *A. chinensis* var. *chinensis* 'Jinpai', two identical *NLR25-1* genes were detected in haplotype 1. Further analysis revealed an asymmetric recombination wherein a large fragment containing an entire *NLR25-1* locus was probably translocated from haplotype 2 into haplotype 1 (Additional file 2: Fig. S26a). A Copia LTR resides in the junction point of the recombination, which probably contributed to the translocation (Additional file 2: Fig. S26b). Intriguingly, when validating this recombination via PCR amplification across the junction point in various *A. chinensis* var. *chinensis* accessions, an unexpected band appeared for *A. chinensis* var. *chinensis* cv. 'Biyu' (Additional file 2: Fig. S26c). Further analysis revealed a similar junction point in haplotype 2 of 'Biyu', suggesting a symmetric recombination might result in an inversion (Additional file 2: Fig. S26a and S26d). As the same junction point appeared in different types of structure variations, it could be assumed that this Copia LTR is relatively active. This may suggest a possible causal agent for the expansion or contraction of pathogen-resistant genes, while also enlightening an approach to enhance resistance utilizing natural source of duplicated pathogen-resistant gene loci.

Taken together, our work demonstrates that there are a large number of TE-mediated structural variations and *cis*-regulatory elements useful for resistance to canker disease in wild kiwifruit germplasm resources, highlighting that once these structural variations are transferred from wild kiwifruit to cultivated varieties, their *cis*-regulatory elements can be activated and utilized by trans-factors derived from both wild and cultivated ones.

Materials and methods

Genome sequencing and assembly

To comprehensively represent the genetic diversity of kiwifruit, a total of eleven samples of *A. eriantha* representing its natural distribution areas and an interspecific hybrid F1

individual (named 'HH-1') between an *A. eriantha* accession 'White' [33] and an *A. chinensis* var. *chinensis* accession 'H0809' [35] were selected for de novo genome assembly. The 12 newly sequenced plants were grown at the farm of Anhui Agricultural University. Fresh and healthy leaves were harvested upon sprouting, promptly frozen in liquid nitrogen, and subsequently used to construct HiFi SMRTbell libraries. The libraries were then sequenced on the PacBio Revio platform using the Circular Consensus Sequencing (CCS) model. Then the HiFi reads of each accession were assembled to contigs of two haplotypes using hifiasm (v0.16.1) [36] with the default parameters. Employing the quarTeT (v1.2.4) [37], the contigs were oriented and anchored to chromosomes by the reference-guided strategy using the "HY4P" (*A. chinensis* var. *chinensis*) and "MDHAPA" (*A. eriantha*) genomes as references [22, 31], respectively. Then the completeness of assemblies was evaluated using BUSCO (v3.0.2) [64] with the Embryophyta odb10 dataset. The continuity was assessed using the LTR Assembly Index (LAI) [38].

Genome annotation

The transposable elements (TEs) were annotated using the Extensive De Novo TE Annotator (EDTA) pipeline [39] with parameters of "–anno 1 –force 1". The protein-coding genes were annotated using the BRAKER3 pipeline [65] through a combination of de novo prediction, homology prediction and transcript evidence. Only those genes that contained both start and stop codons and were at least 100 amino acids (aa) in length were reserved. The function of all protein-coding genes was annotated by eggNOG-mapper [66] against a series of protein sequence databases to obtain the functional description, Gene Ontology (GO) numbers, and PFAM domains.

Gene family-based pangenome construction

By integrating five previously published kiwifruit assemblies, we constructed a pangenome of *A. eriantha* comprising 28 assemblies derived from 15 accessions (Additional file 1: Table S4). Ortholog groups among the 28 assemblies were identified using OrthoFinder [46] with default parameters (-M msa -S diamond -A mafft -T fasttree -I 1.2). Based on the clustering analysis, the resulting families were divided into core (those gene families shared across all 28 assemblies), softcore (those presence > 90% of all accessions, 26–27), dispensable (those present in 2 to 25 assemblies) and private gene families (present in only one assembly). Similarly, all genes from the 46 assemblies were clustered into gene families using the OrthoFinder software [46]. The resulting families were divided into core (those shared across all 46 assemblies), softcore (those presence > 90% of all assemblies, 42–45), dispensable (those shared in 2 to 41 assemblies) and private gene families (those existing in only one assembly).

Detection of genomic variations

A graph-based, reference-unbiased pangenome was constructed using 44 haplotype genomes derived from 22 accessions (comprising thirteen *A. eriantha*, eight *A. chinensis* var. *chinensis*, and one F1 hybrid). The process involved extracting chromosome sequences from all assemblies and renaming them in accordance with PanSN specification rules (e.g., "MD31#1#Chr02" denotes Chromosome 2 on Haplotype 1 of the 'MD31' accession). Following the Pangenome Graph Builder (PGGB v0.7.4) pipeline [43], we initially partitioned

all chromosomal sequences into 29 communities, wherein each community represented one chromosome, and graph-based pangenomes were constructed for each chromosome. The graph construction process followed a standard pipeline: aligning assemblies with wfmash (v0.14.0); generating an initial graph from alignments using seqwish (v0.7.11); refining the graph with smoothxg (v0.8.2) and gfaffix (v0.2.1); finally analyzing and visualizing with odgi (v0.9.2). The key parameters for constructing the graph pangenome were set as "-s 50,000 -p 90 -k 23 -n 44", and the variants were decomposing using MD31.hap1 as reference (-V "MD31#1:1000"). The odgi stats command were used to produce the final statistics of the graph. Based on the VCF file generated from the PGGB, we categorized all variants into three types: SNPs, InDels (< 50 bp), and SVs (≥ 50 bp). Then, SVs with fully discordant genotypes between *A. chinensis* var. *chinensis* and *A. eriantha* accessions, termed highly differentiated SVs, were selected for downstream analysis. The annotated SVs' impact was detected using the SnpEff software [67].

Transcriptome sequencing and data analysis

To evaluate the impact of structural variations (SVs) on gene expression, we performed transcriptome sequencing using young leaves from 16 accessions, including eight *A. eriantha* accessions (MD31, CY38, DG01, GZ26, HN01, MQ32, QS41, YX02) and eight *A. chinensis* var. *chinensis* accessions (Biyu, Donghong, Hongyang, Hort16A, Huangyang, Jinmi, Jinpai, Zps). Three biological replicates were sequenced for each sample. Total RNA was extracted using TIANGEN RNAprep Pure Kit. And then mRNA libraries were constructed and sequenced using the NovoSeq 6000 platform of illumina. The clean reads were mapped to 'MDHAPA' genome [31] using the hisat2 [68] with default parameter. Transcripts per million (TPM) values were calculated using featureCounts [69]. The differential expression analysis was performed using the DESeq2 package [70] in R and the screening criteria was set as '$|\log_2(\text{fold change})| > 1$ and the p-value < 0.05'.

Pan-NLRome construction and analysis

Genes containing the NB-ARC domain were identified by pfamscan (https://github.com/aziele/pfam_scan) with E-value < 1e-5 as candidate NLRs. Then the candidate NLRs were validated and classified by NLR-annotator [71]. Only genes containing the NB-ARC domain were retained as the final NLR genes. Based on the distribution of NB-ARC, Toll/interleukin-1 receptor (TIR), coiled-coil (CC), and leucine-rich repeat (LRR) domains within each protein, these members were classified into six subgroups including TN (TIR and NB-ARC), CN (CC and NB-ARC), NL (NB-ARC and LRR), CNL (CC, NB-ARC, and LRR), NB (NB-ARC), and TNL (TIR, NB-ARC, and LRR). The number of members in the six NLR subgroups was then compared between *A. chinensis* var. *chinensis* and *A. eriantha*. We clustered NLR genes from all 46 assemblies using OrthoFinder [46] with default parameters (-M msa -S diamond -A mafft -T fasttree -I 1.2) to obtain non-redundant NLR orthogroups. Based on the distribution of each non-redundant NLR orthogroup across all assemblies, they were categorized into core (those shared in all 45 assemblies), soft-core (those presence > 90% of all assemblies, 41–45), dispensable (those shared in 2 to 40 assemblies), and private (those existing in only one assembly) NLR orthogroups. The presence-absence patterns of each non-redundant NLR orthogroup across the 46 assemblies were visualized using a heatmap. Using the previously sequenced leaf transcriptomes from 16

accessions, we aligned the reads to their respective reference genomes and quantified NLR gene expression levels. Expression levels of the 99 non-redundant NLR orthogroups were calculated as the mean TPM value of all NLR genes within each orthogroup.

Transcriptome sequencing and allele-specific expression (ASE) analysis

For *Pseudomonas syringae* pv. *actinidiae* (*Psa*) inoculation, five-leaf-stage kiwifruit seedlings were needle-punctured and immersed in bacterial suspension ($OD_{600} = 1.0$, 0.025% Silwet L-77) for 10 min. The inoculated seedlings were maintained in a controlled-environment chamber with the following parameters: 16/8 h photoperiod, 18 °C/14 °C day/night temperature regime, and 95% relative humidity. Leaves of 'HH-1', 'White' and 'H0809' were collected at 0-, 12- and 72-hour-post-infection (hpi) for transcriptome sequencing with at least three biological replicates. All samples were frozen in liquid nitrogen and stored at −80℃. Then total RNA was isolated and purified from sampled tissues and sequenced by Illumina NovoSeq 6000 platform. The RNA-seq raw reads retrieved from 'HH-1' samples (0, 12 and 72 hpi) were mapped against allele-aware annotated gene models according to previous procedure [40], and that from 'White' or 'H0809' samples were aligned to HH-1.hap1 or HH-1.hap2 gene models respectively using hisat2 (v2.1.0) software [68]. The alignment was counted by featureCounts (v2.0.6) software [69] with additional option (-M –fraction) to infer allele-aware expression. More specifically, the reads that contain haplotype-specific SNP were assigned to corresponding haplotype, and reads without SNP were averagely distributed to two haplotypes. The read counts were normalized by the sequencing depth and the transcript length to transcript per kilobase of exon model per million mapped reads (TPM).

The pathogen-resistant genes were annotated though DRAGO and HMMER [54, 55], and the allelic gene pairs were identified utilizing 'jcvi.compara.catalog ortholog' in JCVI library [53]. The transcription factors were identified via PlantTFDB [72]. The differentially expressed genes (DEGs) between 0 and 12 hpi or 0 hpi and 72 hpi sample for 'HH-1', 'White' or 'H0809', the allele-specific expression genes (ASEGs) in 0 hpi, 12 hpi or 72 hpi sample of 'HH-1', as well as DEGs between 'White' and 'H0809' at 0 hpi, 12 hpi or 72 hpi sample were identified using DESeq2 (v1.42.0) software [70] with a standard of $|\ \mathrm{Log_2FoldChange}\ | > 0.5$, and adjusted P-value < 1e-5. HH-1.hap1 was used as reference when conducting ortholog comparison. Transcription factors enrichment was conducted utilizing clusterProfiler package [73]. Co-expression transcription factors of up-regulated genes were filtered with a standard of Pearson correlation coefficient > 0.5.

RNA/DNA extraction and qRT-PCR

The total RNAs were isolated from kiwifruit leaves using the Plant Total RNA Isolation Kit (Vazyme, RC411) following the manufacturer's protocol. First-strand cDNA was synthesized using an All-in-One First-Strand Synthesis MasterMix (with dsDNase) (EG15133S, BestEnzymes Biotech). The CTAB method was used to extract high-quality genomic DNA from the young kiwifruit leaves. Taq SYBR® Green qPCR Premix (Universal) (EG20117M, BestEnzymes Biotech) was used for qRT-PCR on a CFX Connect real-time PCR Detection system (Bio-Rad), the ACTIN was used as reference genes for calibration of qRT-PCR data. The primers of qRT-PCR are included in Additional file 1: Table S33.

Genetic transformation of kiwifruit

For gene knocking out, the editing sites for the target genes were designed through the CRISPOR web tool [74]. *Agrobacterium*-mediated genetic transformation of kiwifruit was performed following the previously described procedure [58]. Briefly, the complete PTG structure of U6-29-tRNA-sgRNA containing targeting sequences was synthesized through gene synthesis service (Sangon Biotech, shanghai, China). The constructed plasmid was transferred into *A. rhizogenes* K599. The buds from 'HH-1' were completely submerged in inducted bacterial solution for 10–12 min, then transferred to a co-cultivation medium for 2 days in the dark. Until the seedlings grew, the induced hairy roots were planted in media containing Timentin. DNA amplification was conducted on transgenic plants to check out positive plants.

For gene overexpression, the complete coding sequence of target gene was amplified from cDNA of resistant maternal plant 'White', and integrated into pBI121 vector. The constructed plasmid was transferred into *A. rhizogenes* K599, and the subsequent operation was performed same as described above.

Psa inoculation and evaluation

Shoots were inoculated through an injury length 1 mm with *Psa* suspension culture of OD_{600} values 0.1 and phenotypic evaluation was performed at 20 days-post-infection (dpi). Leaf discs were inoculated by vacuum osmosis [75] with *Psa* suspension culture with OD_{600} values 0.01 and phenotypic evaluation was performed at 3 dpi. Tissue culture seedlings with uniformly developmental stage were inoculated by flood assay [76] using *Psa* suspension culture with OD_{600} values 0.6 for 15 min and phenotypic evaluation was conducted at 7 dpi. Potted seedlings were inoculated by flood assay using *Psa* suspension culture with OD_{600} value 0.6 for 15 min and the phenotypic evaluation was conducted at 7 dpi.

Physiological indexes measurement

Trypan blue staining (C0040), NBT staining (G4816), DAB staining (G4815) was measured using corresponding assay kits (Beijing Solarbio Science & technology Co., Ltd, Beijing, China) following the manufacturer's instructions. POD activity (AKAO005C), SOD activity (AKAO001C), CAT activity (AKAO003), MDA content (AKFA013C), H_2O_2 content (AKAO009C) and O_2^- content (AKAO008C) were measured using corresponding assay kits (Beijing Boxbio Science & technology Co., Ltd, Beijing, China) following the manufacturer's instructions. The *in planta* bacterial growth was measured by counting the number of bacterial colonies on leaf discs using the spread-plate method [77]. GraphPad Prism 9.5 was used for statistical analyses and visualization.

DNA–protein interaction assays

For dual-luciferase assay, the promoters of *AeNLR25-1* or *AcNLR25-1* were cloned into the pGreen0800-LUC vector and the ORFs of *AeWRKY75* or *AcWRKY75* were inserted into the pBI121 vector, respectively. The corresponding constructs of pGreenII 0800-LUC and pBI121-HA vector were co-injected into *Nicotiana benthamiana* leaves after transforming into *Agrobacterium tumefaciens* strain GV3101. Luminescence was detected 48 hpi (hour-post-infiltration) using 1 mM luciferin with a Tanon 4600SF

system. LUC/REN activity was measured using corresponding assay kits purchased from Yeasen Biotechnology (Shanghai, China) following the manufacturer's instruction.

To construct the plasmid for Yeast-one-hybrid, the full length of *AeWRKY75* or *AcWRKY75* was amplified from the cDNA of 'White' and 'H0809' and cloned into pJG4-5 vector. The promoters of *AeNLR25-1* or *AcNLR25-1* were amplified from the genomic DNA of 'White' and 'H0809' and cloned into pLacZi vector. The corresponding plasmids were co-transformed into the yeast strain EGY48 and were cultured on the SD/-Ura/-Trp selection medium for 48 h, then monoclonal colonies were transferred to the SD/-Ura/-Trp/+ X-gal selection medium for 24 h.

EMSA was conducted using purified AeWRKY75/AcWRKY75 protein and 3' biotin-labeled probes containing the putative WRKY-binding *cis*-element (TGAC/GTCA) from the *AeNLR25-1* promoter. The full-length CDS of *AeWRKY75*/*AcWRKY75* was inserted into the pMal-c2X vector for the MBP fusion protein expression. The fusion plasmid was transformed into Rosetta *Escherichia coli* cells for AeWRKY75/AcWRKY75 protein expression. The probes were 3' biotin end labeled and converted to dsDNA probes by annealing complementary oligonucleotides (Sangon Biotech.) The experiment was performed using the LightShift Chemiluminescent EMSA kit (Beyotime GS009). The binding specificity was examined by competition probe. For mutant probe, the *cis*-element (TGAC) was replaced with TTTT. The primers used were listed in Additional file 1: Table S33.

Agrobacterium-mediated transient expression in kiwifruit

To analyze the promoter activity of the *A. eriantha* 'White' and *A. chinensis* var. *chinensis* 'H0809', the promoter sequence of *NLR25-1* was amplified from the DNA of 'White' and 'H0809' and cloned into the pCAMBIA-GUS vector. Vacuum infiltration was used to perform transient expression in kiwifruit plants. Briefly, seedlings separated from 'HH-1' were infiltrated with suspension culture of *Agrobacterium* strain GV3101 ($OD_{600}=0.8$) with infection buffer (10 mM $MgCl_2 \cdot 6H_2O$) for 10 min. Before infiltrating, Silwet L-77 was added to buffer with a concentration of 0.025%.

The GUS staining was used to detect promoter activity. Leaves from transiently expressing kiwifruit plants were vacuum immersed in staining buffer for 30 min and then incubated overnight at 37 °C. The stained leaves were incubated in 70% ethanol to remove chlorophyll.

Statistical analysis

All the experiments and measurements have been independently performed at least three times. One-way ANOVA (analysis of variance) was used for statistical analysis, and statistical differences were indicated by the different letters ($p<0.05$, Tukey's test). *p-value* for Student's t-test is presented by asterisk (**: $p<0.01$; ***: $p<0.001$).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-026-04068-0.

Additional file 1: This file contains Tables S1-S33.

Additional file 2: This file contains Figures S1-S26.

Additional file 3. This file contains full-length, uncropped scans of all gels and blots.

Peer review information
Martin Mascher and Wenjing She were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Authors' contributions
Y. Liu, S. Wang and J. Yue planned and designed the research. Y. Wu, Y. Wang and Y. Lin conducted experiments and contributed equally to this work. Y. Wang, Y. Lin and Y. Wu analyzed data. M. Zhao, Y. Huo, T. Zhang, H. Wang and Y. Zhu participated in performing experiments. Y. Wang, Y. Wu, Y. Lin and Y. Liu wrote the manuscript. L. Wang, P. Zheng, J. Yue, S. Wang, Y. Wang and Y. Liu revised the manuscript. All authors read and approved the final manuscript.

Funding
This research was supported by the National Natural Science Foundation of China (Grant Nos. U23A20204, 32472680 and 32472703) and Lishui City's Key R&D Program Project (2025zdyf14).

Data availability
The genome sequencing data and assemblies have been deposited in NGDC (https://ngdc.cncb.ac.cn/) and the accession number is PRJCA046593 [78]. Genome assembly and annotation files can also be accessed in Figshare [79]. The variation data of reference-unbiased interspecific graph pangenome has also been deposited in Figshare [80]. The collected materials in the study are available from Yongsheng Liu (liuyongsheng1122@ahau.edu.cn) upon request. The custom scripts and code used in the study are available in a GitHub repository under the MIT License [81] and in a Zenodo repository [82]. Additionally, we utilized publicly available genome assembly of *A. eriantha* 'Blank' [35], *A. eriantha* 'MD31' [31], *A. eriantha* 'Wild' [41], *A. chinensis* var. *chinensis* 'Biyu' [20], *A. chinensis* var. *chinensis* 'Donghong' [23], *A. chinensis* var. *chinensis* 'H0809' [35], *A. chinensis* var. *chinensis* 'Hongyang' [22], *A. chinensis* var. *chinensis* 'Hort16A' [20], *A. chinensis* var. *chinensis* 'Huangyang' [20], *A. chinensis* var. *chinensis* 'Jinmi' [20], *A. chinensis* var. *chinensis* 'Jinpai' [20] and *A. chinensis* var. *chinensis* 'Zps' [20], downloaded from Kiwifruit Pan-Genome Database (KPGD) [62].

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 26 September 2025 Accepted: 2 April 2026
Published online: 09 April 2026

References

1. Ferguson AR. 1904-the year that kiwifruit (*Actinidia deliciosa*) came to New Zealand. N Z J Crop Hortic. 2004;32(1):3–27.
2. Vanneste JL. The scientific, economic, and social impacts of the New Zealand outbreak of bacterial canker of kiwifruit (*Pseudomonas syringae* pv. *actinidiae*). Annu Rev Phytopathol. 2017;55(1):377–99.
3. Wang F, Gao J, Li J, Liu C, Mo Q, Liu P, et al. In vitro evaluation of *Actinidia chinensis* cultivars for their resistance to *Pseudomonas syringae* pv. *actinidiae*. Sci Hortic. 2023. https://doi.org/10.1016/j.scienta.2023.111896.
4. Serizawa S, Ichikawa T, Takikawa Y, Tsuyumu S, Goto M. Occurrence of bacterial canker of kiwifruit in Japan description of symptoms, isolation of the pathogen and screening of bactericides. Jpn J Phytopathol. 1989;55:427–36.
5. Vanneste JL, Poliakoff F, Audusseau C, Cornish DA, Paillard S, Rivoal C, et al. First report of *Pseudomonas syringae* pv. *actinidiae*, the causal agent of bacterial canker of kiwifruit in France. Plant Dis. 2011;95(10):1311.
6. Everett KR, Taylor RK, Romberg MK, Rees-George J, Fullerton RA, Vanneste JL, et al. First report of *Pseudomonas syringae* pv. *actinidiae* causing kiwifruit bacterial canker in New Zealand. Australas Plant Dis Notes. 2011;6(1):67–71.
7. Scortichini M. Occurrence of *Pseudomonas syringae* pv. *actinidiae* on kiwifruit in Italy. Plant Pathol. 1994. https://doi.org/10.1111/j.1365-3059.1994.tb01654.x.
8. Koh YJ: Canker of kiwifruit by *Pseudomonas syringae* pv. *morsprunorum*. Korean J Plant Pathol. 1992;8(2):119-22.
9. Li L, Zhong C, Li D, Zhang S, Huang H. Research progress on bacterial canker disease of kiwifruit. J Huazhong Agric Univ. 2013;32(5):124–33.
10. Li R, Zhang Y, Du X, Wang X, Liu W, Huang L. Calcium-sensing receptor AcCaS regulates chloroplast immunity in kiwifruit by competitively binding with Ca^{2+} or the *Psa* effector. Hortic Res. 2025. https://doi.org/10.1093/hr/uhaf230.
11. Donati I, Buriani G, Cellini A, Mauri S, Costa G, Spinelli F. New insights on the bacterial canker of kiwifruit (*Pseudomonas syringae* pv. *actinidiae*). J Berry Res. 2014;4(2):53–67.
12. Liu Y, Li D, Zhang Q, Song C, Zhong C, Zhang X, et al. Rapid radiations of both kiwifruit hybrid lineages and their parents shed light on a two-layer mode of species diversification. New Phytol. 2017;215(2):877–90.
13. Liu X, Bulley SM, Varkonyi-Gasic E, Zhong C, Li D. Kiwifruit bZIP transcription factor AcePosF21 elicits ascorbic acid biosynthesis during cold stress. Plant Physiol. 2023;192(2):982–99.

14. Liu X, Wu R, Bulley SM, Zhong C, Li D. Kiwifruit MYBS1-like and GBF3 transcription factors influence l-ascorbic acid biosynthesis by activating transcription of GDP-L-galactose phosphorylase 3. New Phytol. 2022;234(5):1782–800.
15. Yuan X, Zheng H, Fan J, Liu F, Li J, Zhong C, et al. Comparative study on physicochemical and nutritional qualities of kiwifruit varieties. Foods. 2022. https://doi.org/10.3390/foods12010108.
16. Hu X, Li L, Wang X, Xiao H, Zhu X, Zhu J, et al. Transcriptional and metabolomic analysis reveal different aromatic amino acid production and potential regulatory mechanisms for two kiwifruit species. Fruit Res. 2025;5(1):0–0.
17. Gan Z, Yupei Z, Caining Y, Qing C, Liqin Z, Wenbin K, et al. A kiwifruit bHLH149 transcription factor modulates carotenoid biosynthesis by directly activating LCYB during postharvest ripening. Fruit Res. 2025;5(1):1–8.
18. Zhao L, Liu Y, Wang M, Xiang L, Wang H, Chen X, et al. MdWRKY20-MdPR1 module mediates resistance of apple to *Fusarium solani*. Fruit Res. 2025;5(1):1–10.
19. Liu Y, Zhou Y, Cheng F, Zhou R, Yang Y, Wang Y, et al. Chromosome-level genome of putative autohexaploid *Actinidia deliciosa* provides insights into polyploidisation and evolution. Plant J. 2024;118(1):73–89.
20. Wang Y, Li P, Zhu Y, Zhang F, Zhang S, He Y, et al. Graph-based pangenome of *Actinidia chinensis* reveals structural variations mediating fruit degreening. Adv Sci. 2024;11(28):e2400322.
21. Zhang F, Wang Y, Lin Y, Wang H, Wu Y, Ren W, et al. Haplotype-resolved genome assembly provides insights into evolutionary history of the *Actinidia arguta* tetraploid. Mol Hortic. 2024;4(1):4.
22. Yue J, Chen Q, Wang Y, Zhang L, Ye C, Wang X, et al. Telomere-to-telomere and gap-free reference genome assembly of the kiwifruit *Actinidia chinensis*. Hortic Res. 2023;10(2):uhac264.
23. Han X, Zhang Y, Zhang Q, Ma N, Liu X, Tao W, et al. Two haplotype-resolved, gap-free genome assemblies for *Actinidia latifolia* and *Actinidia chinensis* shed light on the regulatory mechanisms of vitamin C and sucrose metabolism in kiwifruit. Mol Plant. 2023;16(2):452–70.
24. Alonge M, Wang X, Benoit M, Soyk S, Pereira L, Zhang L, et al. Major impacts of widespread structural variation on gene expression and crop improvement in tomato. Cell. 2020;182(1):145–61.
25. Yu X, Qu M, Wu P, Zhou M, Lai E, Liu H, et al. Super pan-genome reveals extensive genomic variations associated with phenotypic divergence in *Actinidia*. Mol Hortic. 2025;5(1):4.
26. Wu Y, Liu J, Sheng X, Wang W, Wang T, Martinez-Sanchez M, et al. Spatial regulation of chlorophyll degradation in kiwifruit: AcNAC2-AcSGR1/2 cascades mediate rapid de-greening in the inner pericarp. Plant Biotechnol J. 2025;23(7):2554–69.
27. Hu X, Xu C, Li X, Li L, Bao Y, Gu M, et al. Subgenome dominance in allotetraploid *Actinidia valvata* regulates RNA m(6) A modification for waterlogging tolerance. Adv Sci. 2025;12(32):e3974.
28. Testolin R, Huang HW, Ferguson AR. The Kiwifruit Genome, vol. 978-3-319–32274–2. Cham: Springer International Publishing; 2016.
29. Wang T, Ran Y, Atkinson RG, Gleave AP, Cohen D. Transformation of *Actinidia eriantha*: a potential species for functional genomics studies in *Actinidia*. Plant Cell Rep. 2006;25(5):425–31.
30. Wu H, Ma T, Kang M, Ai F, Zhang J, Dong G, et al. A high-quality *Actinidia chinensis* (kiwifruit) genome. Hortic Res. 2019;6:117.
31. Wang Y, Dong M, Wu Y, Zhang F, Ren W, Lin Y, et al. Telomere-to-telomere and haplotype-resolved genome of the kiwifruit *Actinidia eriantha*. Mol Hortic. 2023;3(1):4.
32. Xu XB, Huang CH, Qu XY, Chen M, Zhong M, Lang BB, et al. A new easy peeling *Actinidia eriantha* cultivar "Ganmi 6." Acta Hortic Sin. 2015;42(12):2539–40.
33. Tang W, Sun X, Yue J, Tang X, Jiao C, Yang Y, et al. Chromosome-scale genome assembly of kiwifruit *Actinidia eriantha* with single-molecule sequencing and chromatin interaction mapping. GIGASCIENCE. 2019. https://doi.org/10.1093/gigascience/giz027.
34. Liao G, Huang C, Jia D, Zhong M, Tao J, Qu X, Xu X: High quality genome of *Actinidia eriantha* provide new insight into ascorbic acid regulation. J Integr Agr. 2023;22(11):4244-55.
35. Yue J, Chen Q, Zhang S, Lin Y, Ren W, Li B, et al. Origin and evolution of the kiwifruit Y chromosome. Plant Biotechnol J. 2024;22(2):287–9.
36. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat Methods. 2021;18(2):170–5.
37. Lin Y, Ye C, Li X, Chen Q, Wu Y, Zhang F, et al. Quartet: a telomere-to-telomere toolkit for gap-free genome assembly and centromeric repeat identification. Hortic Res. 2023;10(8):uhad127.
38. Ou S, Chen J, Jiang N. Assessing genome assembly quality using the LTR assembly index (LAI). Nucleic Acids Res. 2018;46(21):e126.
39. Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, et al. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. Genome Biol. 2019;20(1):275.
40. Wang F, Wang S, Wu Y, Jiang D, Yi Q, Zhang M, et al. Haplotype-resolved genome of a papeda provides insights into the geographical origin and evolution of Citrus. J Integr Plant Biol. 2025;67(2):276–93.
41. Yao X, Wang S, Wang Z, Li D, Jiang Q, Zhang Q, et al. The genome sequencing and comparative analysis of a wild kiwifruit *Actinidia eriantha*. Mol Hortic. 2022;2(1):13.
42. Khan AW, Garg V, Roorkiwal M, Golicz AA, Edwards D, Varshney RK. Super-pangenome by integrating the wild side of a species for accelerated crop improvement. Trends Plant Sci. 2020;25(2):148–58.
43. Garrison E, Guarracino A, Heumos S, Villani F, Bao Z, Tattini L, et al. Building pangenome graphs. Nat Methods. 2024;21(11):2008–12.
44. Feehan JM, Castel B, Bentham AR, Jones JD. Plant nlrs get by with a little help from their friends. Curr Opin Plant Biol. 2020;56:99–108.
45. Jacob F, Vernaldi S, Maekawa T. Evolution and conservation of plant nlr functions. Front Immunol. 2013;4:297.
46. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20(1):238.

47. Cameron A, Sarojini V. *Pseudomonas syringae* pv. *actinidiae*: chemical control, resistance mechanisms and possible alternatives. Plant Pathol. 2014;63(1):1–11.
48. Gao J, Liu C, Li J, Li L, Qi B, Gong H, et al. Limited hypersensitive response but enhanced lignin synthesis leads to *Pseudomonas syringae* pv. *actinidiae* tolerance in *Actinidia eriantha*. Horticulture Advances. 2025. https://doi.org/10.1007/s44281-024-00061-4.
49. Zhou Y, Huang S, Tang W, Wu Z, Sun S, Qiu Y, et al. Genomic variation and host interaction among *Pseudomonas syringae* pv. *actinidiae* strains in *Actinidia chinensis* "Hongyang." Int J Mol Sci. 2022. https://doi.org/10.3390/ijms23179743.
50. Kilian B, Dempewolf H, Guarino L, Werner P, Coyne C, Warburton ML. Crop science special issue: adapting agriculture to climate change: a walk on the wild side. Crop Sci. 2021;61(1):32–6.
51. Wang M, Li J, Qi Z, Long Y, Pei L, Huang X, et al. Genomic innovation and regulatory rewiring during evolution of the cotton genus *Gossypium*. Nat Genet. 2022;54(12):1959–71.
52. Gaur U, Li K, Mei S, Liu G. Research progress in allele-specific expression and its regulatory mechanisms. J Appl Genet. 2013;54(3):271–83.
53. Tang H, Krishnakumar V, Zeng X, Xu Z, Taranto A, Lomas JS, et al. JCVI: a versatile toolkit for comparative genomics analysis. Imeta. 2024;3(4):e211.
54. Calle Garcia J, Guadagno A, Paytuvi-Gallart A, Saera-Vila A, Amoroso CG, D'Esposito D, et al. PRGdb 4.0: an updated database dedicated to genes involved in plant disease resistance process. Nucleic Acids Res. 2022;50(D1):D1483–90.
55. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. Nucleic Acids Res. 2018;46(W1):W200-4.
56. Bournonville CFG, Díaz-Ricci JC. Quantitative determination of superoxide in plant leaves using a modified NBT staining method. Phytochem Anal. 2011;22(3):268–71.
57. Li Y, Zhang D, Wang X, Bai F, Li R, Zhou R, et al. LACCASE35 enhances lignification and resistance against *Pseudomonas syringae* pv. *actinidiae* infection in kiwifruit. Plant Physiol. 2025. https://doi.org/10.1093/plphys/kiaf040.
58. Li P, Zhang Y, Liang J, Hu X, He Y, Miao T, et al. *Agrobacterium rhizogenes*-mediated marker-free transformation and gene editing system revealed that AeCBL3 mediates the formation of calcium oxalate crystal in kiwifruit. Mol Hortic. 2024;4(1):1.
59. Allaby RG, Stevens CJ, Kistler L, Fuller DQ. Emerging evidence of plant domestication as a landscape-level process. Trends Ecol Evol. 2022;37(3):268–79.
60. Trucchi E, Benazzo A, Lari M, Iob A, Vai S, Nanni L, et al. Ancient genomes reveal early Andean farmers selected common beans while preserving diversity. Nat Plants. 2021;7(2):123–8.
61. Purugganan MD. Evolutionary insights into the nature of plant domestication. Curr Biol. 2019;29(14):R705-14.
62. Li B, Li X, Wang Y, Liu X, Li K, Li R, et al. KPGD: a kiwifruit pangenome database for comprehensive mining of genetic diversity in the genus *Actinidia*. Plant Commun. 2025;6(8):101373.
63. Yu X, Qin M, Qu M, Jiang Q, Guo S, Chen Z, et al. Genomic analyses reveal dead-end hybridization between two deeply divergent kiwifruit species rather than homoploid hybrid speciation. Plant J. 2023;115(6):1528–43.
64. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210–2.
65. Gabriel L, Bruna T, Hoff KJ, Ebel M, Lomsadze A, Borodovsky M, et al. BRAKER3: fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. Genome Res. 2024;34(5):769–77.
66. Cantalapiedra CP, Hernandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021;38(12):5825–9.
67. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly. 2012;6(2):80–92.
68. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15.
69. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.
70. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
71. Steuernagel B, Witek K, Krattinger SG, Ramirez-Gonzalez RH, Schoonbeek H, Yu G, et al. The nlr-annotator tool enables annotation of the intracellular immune receptor repertoire. Plant Physiol. 2020;183(2):468–82.
72. Jin J, Tian F, Yang D, Meng Y, Kong L, Luo J, et al. Planttfdb 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. 2017;45(D1):D1040–5.
73. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. Clusterprofiler 4.0: a universal enrichment tool for interpreting omics data. The Innovation. 2021;2(3):100141.
74. Concordet J, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res. 2018;46(W1):W242-5.
75. Zhao C, Liu W, Zhang Y, Li Y, Ma C, Tian R, et al. Two transcription factors, AcREM14 and AcC3H1, enhance the resistance of kiwifruit *Actinidia chinensis* var. *chinensis* to *Pseudomonas syringae* pv. *actinidiae*. Hortic Res. 2024;11(1):uhad242.
76. Ishiga T, Sakata N, Nguyen VT, Ishiga Y. Flood inoculation of seedlings on culture medium to study interactions between *Pseudomonas syringae* pv. *actinidiae* and kiwifruit. J Gen Plant Pathol. 2020;86(4):257–65.
77. Yuan M, Xin X: Bacterial Infection and Hypersensitive Response Assays in *Arabidopsis-Pseudomonas syringae* Pathosystem. Bio-Protocol. 2021;11(24):e4268.
78. Wang Y. The super pangenome of kiwifruit. PRJCA046593. National Genomics Data Center. 2025. https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA046593.
79. Wang Y, Wu Y, Lin Y. Chromosome-scale genome assemblies and annotation for eleven wild *Actinidia eriantha* accessions and one interspecific hybrid between *Actinidia eriantha* and cultivated *Actinidia chinensis* var. *chinensis*. Figshare. 2026. https://doi.org/10.6084/m9.figshare.31884859.

80. Wang Y. Reference-unbiased interspecific graph pangenome of *A. eriantha* and *A. chinensis*. Figshare. 2026. https://doi.org/10.6084/m9.figshare.31004944.
81. Lin Y, Wang Y. Archive of scripts used in publications. GitHub. 2026. https://github.com/Echoring/ScriptArchive/tree/pan-F1.
82. Lin Y, Wang Y. Custom script used in pan/F1 project. 2025. Zenodo. https://doi.org/10.5281/zenodo.18083514.

Publisher's Note