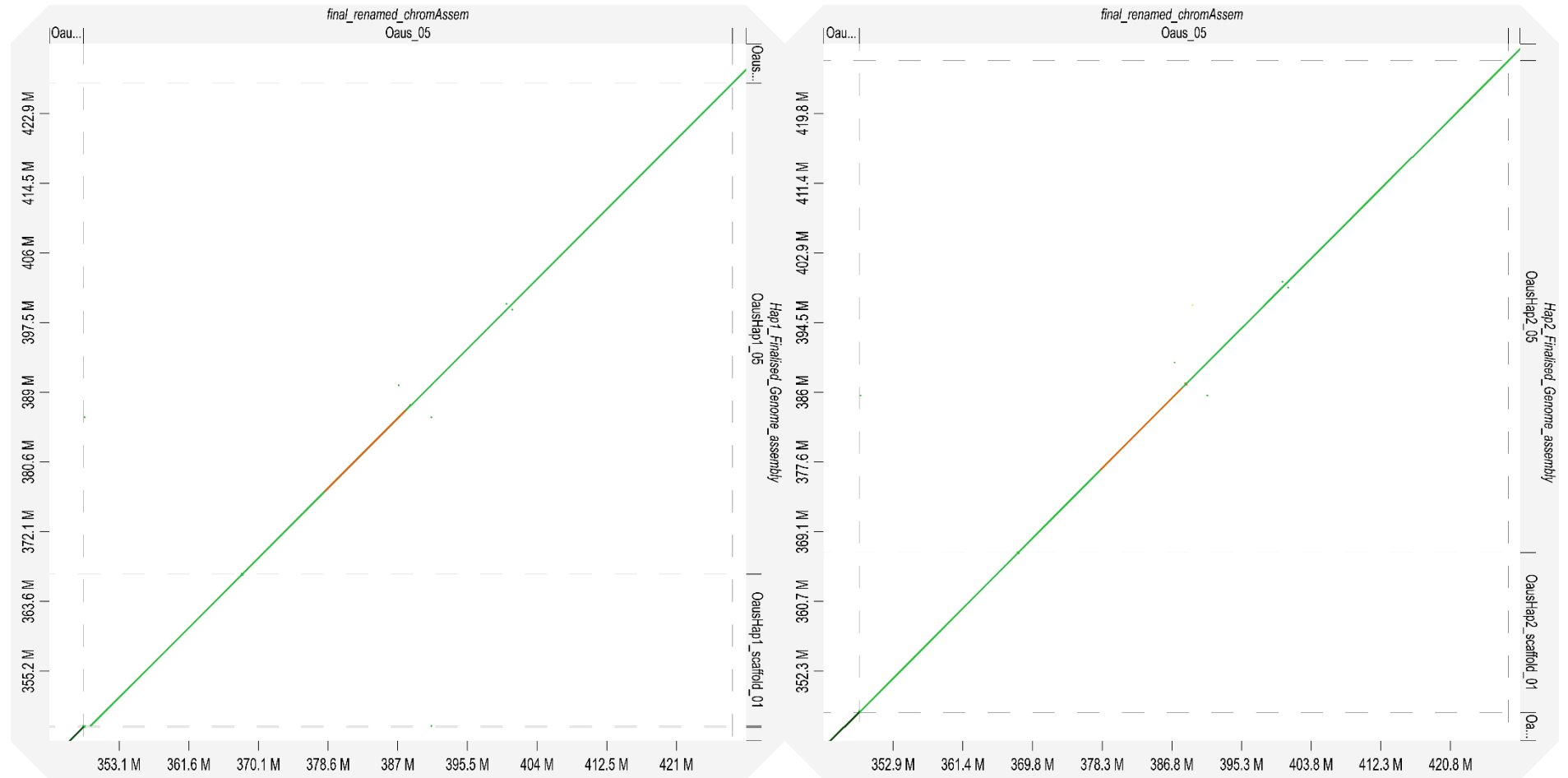
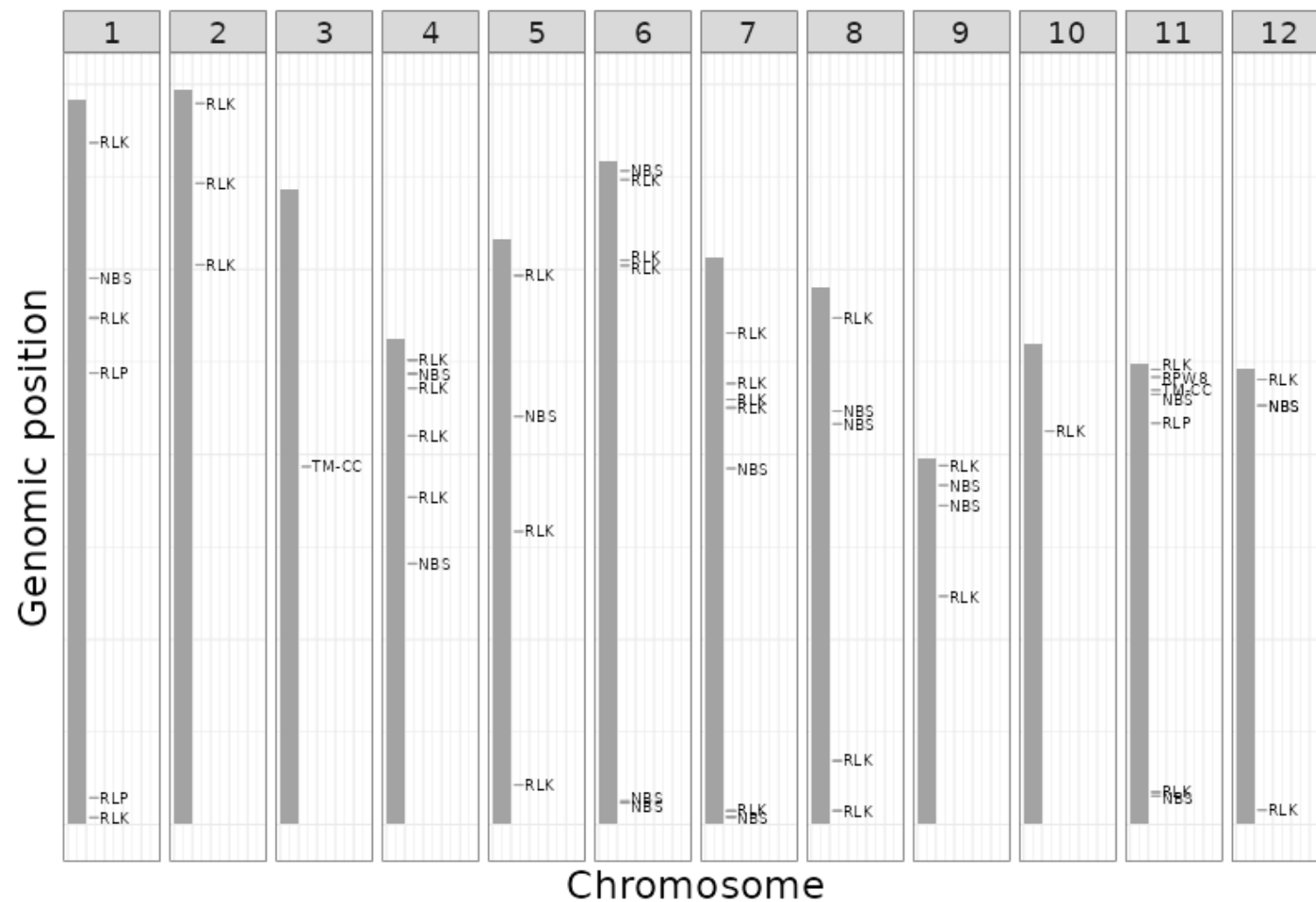


Supplementary Figures



Supplementary figure 1. Chromosome 5 in the collapsed assembly aligns with two scaffolds in both the Hap1 (left) and Hap2 (right) assemblies.



Supplementary figure 2. Distribution of tandem duplicated resistance gene analogue (RGA) clusters and corresponding RGA class type across *O. australiensis* chromosomes. RLK = receptor-like kinase, RLP = receptor-like protein, NBS = nucleotide binding-site-encoding, TM-CC = transmembrane-coiled-coil, RPW8 = resistance to powdery mildew 8 domain.

Supplementary Tables

Supplementary table 1. Finalised scaffolds for primary and haplotype *O. australiensis* assemblies.

Oaus_01	97,953,927	2	OausHap1_01	97,817,653	2	OausHap2_01	97,842,899	2
Oaus_02	99,334,057	2	OausHap1_02	99,281,607	2	OausHap2_02	98,984,626	2
Oaus_03	85,917,448	2	OausHap1_03	85,801,739	2	OausHap2_03	84,041,375	1
Oaus_04	65,507,800	2	OausHap1_04	65,432,925	2	OausHap2_04	64,819,042	2
Oaus_05	79,155,724	2	OausHap1_05	59,677,988	1	OausHap2_05	59,591,565	1
Oaus_06	89,579,249	2	OausHap1_06	89,410,254	2	OausHap2_06	89,387,264	2
Oaus_07	76,570,987	2	OausHap1_07	76,584,184	2	OausHap2_07	76,444,717	2
Oaus_08	72,629,113	2	OausHap1_08	72,406,816	2	OausHap2_08	69,627,118	2
Oaus_09	49,423,460	1	OausHap1_09	49,403,252	1	OausHap2_09	49,382,411	1
Oaus_10	64,960,291	2	OausHap1_10	64,885,109	2	OausHap2_10	63,198,941	2
Oaus_11	62,295,802	2	OausHap1_11	62,183,908	2	OausHap2_11	58,859,920	2
Oaus_12	61,684,835	2	OausHap1_12	61,708,205	2	OausHap2_12	61,466,291	2
Oaus_scaffold_01	3,672,952	0	OausHap1_scaffold_01	18,432,035	0	OausHap2_scaffold_01	19,329,382	1
			OausHap1_scaffold_02	3,204,154	0	OausHap2_scaffold_02	3,240,286	0
			OausHap1_scaffold_03	243,768	1	OausHap2_scaffold_03	2,083,956	0
			OausHap1_scaffold_04	168,424	1	OausHap2_scaffold_04	2,002,091	0
						OausHap2_scaffold_05	1,628,508	0
						OausHap2_scaffold_06	1,584,277	1

Supplementary table 2. Large inversions between current and previous *O. australiensis* assemblies.

Chromosome	<i>O. australiensis</i> inversion starting position	<i>O. australiensis</i> inversion end position	Long et al. (2024) assembly inversion starting position	Long et al. (2024) assembly inversion end position	Inversion size (Mb)*	% of total chromosome length*	ID
Chr1	32557716	34921215	33291222	35652728	2.36	2.41	INV8208
Chr2	22432194	26690570	22407112	26616484	4.26	4.29	INV8255
Chr4	17157621	19142241	16536500	19037977	1.98	3.03	INV8282
Chr4	27113950	28220158	25254099	26424827	1.11	1.69	INV8291
Chr4	23672640	24731814	22293806	23013862	1.06	1.62	INV8286
Chr6	66954358	70055309	65211462	68306854	3.10	3.46	INV8323
Chr6	50203791	52700100	48181227	50732836	2.50	2.79	INV8319
Chr6	29775964	32027437	28955922	30891286	2.25	2.51	INV8312
Chr6	75853572	76976207	73709286	74783092	1.12	1.25	INV8328
Chr7	21979817	28110943	22334571	29657514	6.13	8.01	INV8335
Chr7	13316545	15310663	13567540	15422683	1.99	2.60	INV8334
Chr7	38655831	39850533	39935292	41221648	1.19	1.56	INV8337
Chr8	34325719	38562229	33509323	37643979	4.24	5.83	INV8353
Chr8	38574678	41823887	37817109	40849738	3.25	4.47	INV8354
Chr9	72	1089184	2414638	3546993	1.09	2.20	INV8358
Chr10	13650680	18082845	13770499	17686823	4.43	6.82	INV8221
Chr11	15857307	21466729	13392749	19329754	5.61	9.00	INV8235

*In reference (current *O. australiensis* assembly). Previously published *O. australiensis* assembly from Long et al. 2024. Large inversions >1 Mb and >1.25% of total chromosome length.

Supplementary table 3. GO annotation of species-specific protein clusters in *O. australiensis*.

Cluster ID	Number of proteins	Swiss-Prot Hit	GO Annotation
cluster26	41	Q94HW2: Retrovirus-related Pol polyprotein from transposon RE1	GO:0006310; P:DNA recombination; IEA:UniProtKB-KW
cluster82	23	-	-
cluster94	21	-	-
cluster110	19	Q9M2U3: Protein ALP1-like {ECO:0000305}	GO:0004518; F:nuclease activity; IEA:UniProtKB-KW
cluster153	16	-	-
cluster170	15	-	-
cluster171	15	-	-
cluster227	13	-	-
cluster263	12	P10978: Retrovirus-related Pol polyprotein from transposon TNT 1-94	GO:0015074; P:DNA integration; IEA:InterPro
cluster264	12	P04323: Retrovirus-related Pol polyprotein from transposon 17.6	GO:0015074; P:DNA integration; IEA:InterPro
cluster316	11	A2XNS1: Actin-3	GO:0005524; F:ATP binding; IEA:UniProtKB-KW
cluster317	11	-	-
cluster318	11	Q9SH66: Dirigent protein 5	GO:1901599; P:(-)-pinosresinol biosynthetic process; IDA:TAIR
cluster386	10	Q94HW2: Retrovirus-related Pol polyprotein from transposon RE1	GO:0006310; P:DNA recombination; IEA:UniProtKB-KW
cluster387	10	Q3E7K8: Polyubiquitin 12	GO:0006511; P:ubiquitin-dependent protein catabolic process; ISS:TAIR
cluster504	9	A2XF66: Histone H2B.1	GO:0046982; F:protein heterodimerization activity; IEA:InterPro
cluster505	9	-	-
cluster506	9	-	-
cluster507	9	-	-
cluster638	8	Q43594: Tubulin beta-1 chain	GO:0000278; P:mitotic cell cycle; IBA:GO_Central

cluster639	8	P10978: Retrovirus-related Pol polyprotein from transposon TNT 1-94	GO:0015074; P:DNA integration; IEA:InterPro
cluster640	8	Q75I75: Expansin-A31	GO:0009664; P:plant-type cell wall organization; IEA:InterPro
cluster641	8	-	-
cluster642	8	P43214: Pollen allergen Phl p 2	GO:0005576; C:extracellular region; IEA:UniProtKB-SubCell
cluster643	8	-	-
cluster848	7	-	-
cluster849	7	Q9ZT94: Retrovirus-related Pol polyprotein from transposon RE2	GO:0006310; P:DNA recombination; IEA:UniProtKB-KW
cluster850	7	Q9LHA8: Probable mediator of RNA polymerase II transcription subunit 37c	GO:0009615; P:response to virus; IEP:TAIR
cluster851	7	Q01595: Cortical cell-delineating protein	-
cluster852	7	Q6YZA4: Germin-like protein 8-6	GO:0045735; F:nutrient reservoir activity; IEA:InterPro
cluster853	7	Q4V3D9: Protein STRICTOSIDINE SYNTHASE-LIKE 10 {ECO:0000303 PubMed:10777701}	GO:0009058; P:biosynthetic process; IEA:InterPro
cluster854	7	Q9ZUF1: F-box protein At2g05970	-
cluster1254	6	-	-
cluster1255	6	-	-
cluster1256	6	B7F8P5: Auxin-responsive protein SAUR36	GO:2000012; P:regulation of auxin polar transport; IMP:UniProtKB
cluster1257	6	Q8LIL0: Bisdemethoxycurcumin synthase	GO:0009813; P:flavonoid biosynthetic process; IDA:UniProtKB
cluster1258	6	P35792: Pathogenesis-related protein PRB1-2	GO:0009607; P:response to biotic stimulus; IEA:UniProtKB-KW
cluster1259	6	Q9LXQ8: F-box/LRR-repeat protein At3g58940	-
cluster1260	6	B7F8P5: Auxin-responsive protein SAUR36	GO:2000012; P:regulation of auxin polar transport; IMP:UniProtKB
cluster1889	5	-	-
cluster1890	5	-	-
cluster1891	5	-	-
cluster1892	5	Q06396: ADP-ribosylation factor 1	GO:0016192; P:vesicle-mediated transport; IBA:GO_Central

cluster1893	5	-	-
cluster1894	5	Q03278: Retrovirus-related Pol polyprotein from type-1 retrotransposable element R2	GO:0003964; F:RNA-directed DNA polymerase activity; IEA:UniProtKB-KW
cluster1895	5	-	-
cluster1896	5	-	-
cluster1897	5	P41044: Calbindin-32	GO:0099509; P:regulation of presynaptic cytosolic calcium ion concentration; IBA:GO_Central
cluster1898	5	Q766C3: Aspartic proteinase nepenthesin-1	GO:0004190; F:aspartic-type endopeptidase activity; IEA:UniProtKB-KW
cluster1899	5	Q9LMP1: Wall-associated receptor kinase 2	GO:0009826; P:unidimensional cell growth; IMP:TAIR
cluster3651	4	-	-
cluster3652	4	-	-
cluster3653	4	-	-
cluster3654	4	-	-
cluster3655	4	-	-
cluster3656	4	O94925: Glutaminase kidney isoform, mitochondrial	GO:0001967; P:suckling behavior; IEA:Ensembl
cluster3657	4	Q6K8S5: Putative cyclin-F1-2	GO:0007088; P:regulation of mitotic nuclear division; IBA:GO_Central
cluster3658	4	-	-
cluster3659	4	-	-
cluster3660	4	P09617: Leaf-specific thionin	GO:0006952; P:defense response; IEA:UniProtKB-KW
cluster3661	4	Q7EZT1: NAC domain-containing protein 67 {ECO:0000303 PubMed:15029955}	GO:0006355; P:regulation of transcription, DNA-templated; IEA:InterPro
cluster3662	4	P49211: 60S ribosomal protein L32-1	GO:0006412; P:translation; IEA:InterPro
cluster3663	4	-	-
cluster6302	3	C0LGT6: LRR receptor-like serine/threonine-protein kinase EFR	GO:0010359; P:regulation of anion channel activity; IMP:TAIR
cluster6303	3	-	-
cluster6304	3	-	-
cluster6305	3	Q53RJ5: Heat shock 70 kDa protein BIP2 {ECO:0000305}	GO:0030433; P:ubiquitin-dependent ERAD pathway; IBA:GO_Central
cluster6306	3	-	-

cluster6307	3	-	-
cluster6308	3	-	-
cluster6309	3	-	-
cluster6310	3	-	-
cluster6311	3	Q8RWZ3: Probable acyl-CoA dehydrogenase IBR3 {ECO:0000305}	GO:0048767; P:root hair elongation; IMP:TAIR
cluster6312	3	-	-
cluster6313	3	Q851R5: Thioredoxin H2-2	GO:0006662; P:glycerol ether metabolic process; IEA:InterPro
cluster6314	3	Q0J3D9: Coatomer subunit alpha-3	GO:0006890; P:retrograde vesicle-mediated transport, Golgi to ER; IBA:GO_Central
cluster6315	3	-	-
cluster6316	3	-	-
cluster6317	3	Q10PW9: Sugar transport protein MST4 {ECO:0000305}	GO:0005351; F:carbohydrate:proton symporter activity; IBA:GO_Central
cluster6318	3	Q688S6: Probable metal-nicotianamine transporter YSL4	GO:0055085; P:transmembrane transport; IEA:InterPro
cluster6319	3	-	-
cluster6320	3	-	-
cluster6321	3	Q05968: Pathogenesis-related protein 1	GO:0009607; P:response to biotic stimulus; IEA:UniProtKB-KW
cluster6322	3	Q7XC35: Cyclin-P4-1	GO:0000079; P:regulation of cyclin-dependent protein serine/threonine kinase activity; IEA:InterPro
cluster6323	3	-	-
cluster6324	3	-	-
cluster6325	3	Q2QXJ0: Putative germin-like protein 12-4	GO:0045735; F:nutrient reservoir activity; IEA:InterPro

GO annotation and clustering of homologous proteins performed in OrthoVenn3 (Sun et al., 2023).

Supplementary table 4. Tandem duplicated resistance gene analogue clusters within *O. australiensis*

Chromosome	RGA class	Genomic position	Gene IDs
1	RLK	916515-962965	g83
			g84
	RLP	3600478-3637643	g310
			g311
	RLP	60992424-61004348	g1593
			g1594
	RLK	68368673-68494786	g1889
			g1890
			g1891
			g1893
			g1894
2	NBS	73810671-73826876	g2151
			g2152
	RLK	92092728-92127555	g3054
			g3055
			g5260
			g5261
3	RLK	86631601-86640675	g5661
			g5662
	RLK	97413400-97432108	g6194
4	TM-CC	48349175-48362289	g8289
			g8290
	NBS	35189841-35290601	g9997
			g9998
	RLK	44176953-44262984	g10229
			g10230
	RLK	52524646-52530344	g10838
			g10839
	RLK	58920725-58963409	g11167
			g11168
5	NBS	60886046-60931167	g11339
			g11340
			g11341

4	RLK	62745472-62778503	g11342
			g11484
			g11485
			g11486
			g11487
			g11488
			g11489
5	RLK	5335806-5370371	g12128
			g12129
	RLK	39582150-39601959	g12477
			g12478
	NBS	55127045-55137372	g12817
			g12818
	RLK	74133800-74200454	g13324
		g13327	
		g13328	
6	NBS	2889496-2936106	g14055
			g14056
	NBS	3176892-3186182	g14082
			g14083
	RLK	75462923-75588159	g14962
			g14963
	RLK	76215361-76267842	g14972
		g14973	
7	RLK	87093348-87132411	g15401
			g15402
	NBS	88301304-88318281	g15510
			g15511
	NBS	1014957-1057945	g15718
			g15719
			g15720
		g15722	
		g15723	
		g15724	
	RLK	1786980-1857483	g15781
			g15782
			g15783
			g15784
			g15785
			g15786

7			g15787
	NBS	48105906-48119347	g16418 g16419
	RLK	56219812-56400521	g16584 g16586 g16587 g16588
	RLK	57412044-57428777	g16601 g16602
	RLK	59531398-59644318	g16649 g16650
8	RLK	66384652-66414766	g16769 g16770
	RLK	1832316-1895260	g17513 g17514 g17515
	RLK	8551564-8652916	g17830 g17831 g17832
	NBS	54100347-54135542	g18220 g18221
	NBS	55847732-55856630	g18260 g18261
9	RLK	68441435-68450453	g18622 g18623
	RLK	30837528-30874638	g19197 g19198
	NBS	43064220-43116404	g19675 g19676
	NBS	45836431-45848547	g19886 g19887
10	RLK	48484012-48505878	g20150 g20152
	RLK	53137893-53143293	g20749 g20750
11	NBS	3838381-3881567	g21753 g21754
	RLK	4340751-4388525	g21792

			g21793
			g21794
	RLP	54197668-54251805	g22331
			g22332
	NBS	58107167-58116162	g22501
			g22502
11	TM-CC	58689976-58712969	g22513
			g22516
	RPW8	60444484-60463917	g22594
			g22595
	RLK	61464409-61497587	g22645
			g22646
	RLK	1946864-1957466	g22851
			g22852
			g23591
12	NBS	56592653-56624705	g23592
			g23593
			g23791
	RLK	60111901-60127693	g23792

RLK = receptor-like kinase, RLP = receptor-like protein, NBS = nucleotide binding-site-encoding, TM-CC = transmembrane-coiled-coil, RPW8 = resistance to powdery mildew 8 domain.

Supplementary table 5. Current cloned disease resistance genes in rice from Simon et al. 2023.

Pathogen	Gene	Chromosome	MSU/RAP ID
Brown Planthopper	BPH14	3	Os03g63150
	BPH17	4	Os04g12540– Os04g12560– Os04g12580
	BPH6	4	Os04g35210
	BPH30	4	Os04g02520
	BPH32=BPH3	6	Os06g03240
	BPH29	6	Os06g01860
	BPH1=BPH10=	12	Os12g37290
	BPH18=BPH21/BPH2=		
	/BPH26/BPH9/BPH7		
Bacterial blight/<i>Xanthomonas</i> <i>oryzae</i> pv. <i>oryzae</i>	Xa1/Xa2=Xa31(t)/Xa14/ Xa45(t)/CGS-Xo1	4	Os04g53120
	xa5	5	Os05g01710
	Xa27	6	Os06g39810
	xa13	8	Os08g42350
	Xa3=Xa26	11	Os11g47210
	Xa4	11	Os11g47140
	Xa23	11	Os11g0586701*
	Xa21	11	Os11g36180
	Xa10	11	Os11g37570
	xa41	11	Os11g31190
	Xa46 (t)	11	Os11g0586100*
	Xa47(t)	11	Os11g46200
	xa25	12	Os12g29220
Rice blast/<i>Pyricularia</i> <i>oryzae</i>	Pish/Pi35	1	Os01g57340
	Pit	1	Os01g05620
	Pi37	1	Os01g57310
	Pi64	1	Os01g57280
	Pib	2	Os02g57310
	bsr-d1	3	Os03g32230
	pi21	4	Os04g32850
	Pi63	4	Os04g52970
	Pi9/Pi2/Piz- t/PigmR/Pizh/Pi50	6	Os06g17900
	Pid2	6	Os06g29810
	Pi25/Pid3/Pid3-i1	6	Os06g22460
	Pid4	6	Os06g17950
	Pi36	8	Os08g05440

	Pi5	9	Os09g15840
	Pi56	9	Os09g16000
	Pik/Pik-m/Pik-p/Pi1/Pi7	11	Os11g46200- Os11g46210
	Pia/Pi-CO39	11	Os11g11790- Os11g11810
	Pik-h=Pi54	11	Os11g42010
	NLS-1	11	Os11g0249000*
	Pi65	11	Os11g0694600*
	Pb1	11	Os11g38580
	Pb2	11	Os11g45620
	Pb3	11	Os11g45090
	Pita	12	Os12g18360
	Pita2=Ptr	12	Os12g18729
Rice tungro spherical virus	tsv1	7	Os07g36940
Rice yellow mottle virus	RYMV2	1	Os01g68970
	RYMV1	4	Os04g42140
	RYMV3	11	Os11g43700
Rice stripe virus	STV11	11	Os11g30910

*Rice Annotation Project (RAP) ID. All other IDs are sourced from the MSU Rice Genome Annotation Project. Cloned disease resistance genes originally compiled from Simon et al. 2023². '=' refers to identical alleles, '/' refers to different alleles at the same locus.