

Recombination Rate Heterogeneity within Arabidopsis Disease Resistance Genes

Kyuha Choi, Carsten Reinhard, Heïdi Serra, Piotr A. Ziolkowski, Charles J. Underwood, Xiaohui Zhao, Thomas J. Hardcastle, Nataliya E. Yelina, Catherine Griffin, Matthew Jackson, et al.

▶ **To cite this version:**

Kyuha Choi, Carsten Reinhard, Heïdi Serra, Piotr A. Ziolkowski, Charles J. Underwood, et al.. Recombination Rate Heterogeneity within Arabidopsis Disease Resistance Genes. PLoS Genetics, 2016, 12 (7), pp.1-30. 10.1371/journal.pgen.1006179 . hal-02639877

HAL Id: hal-02639877
https://hal.inrae.fr/hal-02639877

Submitted on 28 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PLOS | GENETICS

RESEARCH ARTICLE

Recombination Rate Heterogeneity within Arabidopsis Disease Resistance Genes

Kyuha Choi[1☯], Carsten Reinhard[1☯], Heïdi Serra[1], Piotr A. Ziolkowski[1,2], Charles J. Underwood[1,3], Xiaohui Zhao[1], Thomas J. Hardcastle[1], Nataliya E. Yelina[1], Catherine Griffin[1], Matthew Jackson[1], Christine Mézard[4], Gil McVean[5], Gregory P. Copenhaver[6,7], Ian R. Henderson[1]*

1 Department of Plant Sciences, Downing Street, University of Cambridge, Cambridge, United Kingdom, **2** Department of Biotechnology, Adam Mickiewicz University, Poznan, Poland, **3** Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States of America, **4** Institut Jean-Pierre Bourgin, INRA, AgroParisTech, CNRS, Université Paris-Saclay, RD10, Versailles, France, **5** The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, **6** Department of Biology and the Integrative Program for Biological and Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America, **7** Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine, Chapel Hill, North Carolina, United States of America

☯ These authors contributed equally to this work.
* irh25@cam.ac.uk

CrossMark
click for updates

OPEN ACCESS

Citation: Choi K, Reinhard C, Serra H, Ziolkowski PA, Underwood CJ, Zhao X, et al. (2016) Recombination Rate Heterogeneity within Arabidopsis Disease Resistance Genes. PLoS Genet 12(7): e1006179. doi:10.1371/journal.pgen.1006179

Editor: Richard Michelmore, UC Davis, UNITED STATES

Received: November 11, 2015

Accepted: June 15, 2016

Published: July 14, 2016

Copyright: © 2016 Choi et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The fastq sequence files for (i) the RAC crossover sequencing experiments and (ii) MNase-seq mapping of nucleosome positions are available from ArrayExpress (E-MTAB-4556). The fastq sequence files for the ColxLer F2 experiments are available from ArrayExpress (E-MTAB-4657). All other data are within the paper and its Supporting Information files.

Funding: Research in the Henderson laboratory was supported by a Royal Society University Research Fellowship, Gatsby Charitable Foundation grant 2962, BBSRC grant BB/N007557/1 and National Natural Science Foundation of China grant

Abstract

Meiotic crossover frequency varies extensively along chromosomes and is typically concentrated in hotspots. As recombination increases genetic diversity, hotspots are predicted to occur at immunity genes, where variation may be beneficial. A major component of plant immunity is recognition of pathogen Avirulence (Avr) effectors by resistance (*R*) genes that encode NBS-LRR domain proteins. Therefore, we sought to test whether NBS-LRR genes would overlap with meiotic crossover hotspots using experimental genetics in *Arabidopsis thaliana*. NBS-LRR genes tend to physically cluster in plant genomes; for example, in Arabidopsis most are located in large clusters on the south arms of chromosomes 1 and 5. We experimentally mapped 1,439 crossovers within these clusters and observed NBS-LRR gene associated hotspots, which were also detected as historical hotspots via analysis of linkage disequilibrium. However, we also observed NBS-LRR gene coldspots, which in some cases correlate with structural heterozygosity. To study recombination at the fine-scale we used high-throughput sequencing to analyze ~1,000 crossovers within the *RESISTANCE TO ALBUGO CANDIDA1* (*RAC1*) *R* gene hotspot. This revealed elevated intragenic crossovers, overlapping nucleosome-occupied exons that encode the TIR, NBS and LRR domains. The highest *RAC1* recombination frequency was promoter-proximal and overlapped CTT-repeat DNA sequence motifs, which have previously been associated with plant crossover hotspots. Additionally, we show a significant influence of natural genetic variation on NBS-LRR cluster recombination rates, using crosses between Arabidopsis ecotypes. In conclusion, we show that a subset of NBS-LRR genes are strong hotspots, whereas others are coldspots. This reveals a complex recombination landscape in Arabidopsis NBS-LRR genes, which we propose results from

61403318. KC was funded by an EMBO long term postdoctoral fellowship ALTF 807-2009. PAZ was supported by a Polish Mobility Plus Fellowship 605/MOB/2011/0. GPC is funded by a National Science Foundation Grant (MCB-1121563). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

varying coevolutionary pressures exerted by host-pathogen relationships, and is influenced by structural heterozygosity.

Author Summary

The majority of plants, animals and fungi reproduce sexually, which has a profound effect on patterns of genetic diversity. For sexual reproduction to occur a specialized cell division called meiosis generates gametes (sex cells) with recombined DNA. During meiosis homologous chromosomes undergo reciprocal exchange of genetic material called crossover, which mixes parental sequences. Crossover frequency is highly variable along chromosomes and is typically concentrated in narrow regions called hotspots. Sexual reproduction is proposed to provide a major advantage during host-pathogen coevolution. This theory predicts that crossover hotspots will be associated with genes that play important roles in the immune system, where diversity may be beneficial. To investigate this idea we mapped patterns of crossover frequency in relation to genes that plants use to recognize pathogens. We observe that a subset of *Arabidopsis thaliana* resistance genes overlap with strong crossover hotspots. However, we also observe resistance gene coldspots, which correlate with chromosomal rearrangements between varietal lines (ecotypes). Using Arabidopsis ecotypes we also show that natural genetic variation has a strong modifying effect on resistance gene cluster recombination rates. Together our work shows complex variation in resistance gene crossover patterns, which may relate to heterogeneity in coevolutionary pressures experienced at different loci, in addition to the effects of structural genetic diversity.

Introduction

During meiosis, homologous chromosomes recombine via reciprocal exchanges of DNA, called crossovers [1,2]. Crossovers initiate from DNA double strand breaks that are repaired using an interhomolog pathway [1–3]. Meiotic recombination frequency is highly variable along chromosomes and is concentrated in narrow (1–2 kb) hotspots in plants, animals and fungi [4–7]. Hotspot locations and activity are controlled by both genetic and epigenetic factors [4–7]. The "two-speed" genome hypothesis proposes that crossover hotspots will be differentially distributed among immunity versus housekeeping genes, as the former participate directly in host-pathogen coevolution, where recombination and diversity may be beneficial [8–12]. Here we investigate this hypothesis by experimentally mapping patterns of meiotic crossover recombination in relation to Arabidopsis NBS-LRR disease resistance genes.

The plant immune system consists of two major branches: (i) Pathogen Associated Molecular Pattern (PAMP) Triggered Immunity (PTI) and (ii) Effector Triggered Immunity (ETI) [13–15]. PTI signaling recognizes conserved PAMPs shared by many pathogenic microbes, including flagellin and chitin, via the action of receptor-like kinases [15]. ETI signaling involves recognition of specific pathogen effectors (the products of *Avirulence* genes, *Avr*) by matching plant disease resistance genes (*R* genes) [13,14]. *R* gene products typically share combinations of domains including TIR (Toll, Interleukin-1 and *R* genes), NBS (Nucleotide Binding Site) and LRR (Leucine Rich Repeats) [13,14,16,17]. R proteins can interact directly with pathogen Avirulence proteins (e.g. RRS1-R & Pop2 and L & AvrL567), or guard host targets and decoys (e.g. RPS2, RIN4 & AvrRpt2) [18–23]. Pathogen recognition by R proteins triggers

hypersensitive cell death, in addition to systemic signaling that increases resistance in uninfected tissues [13,14]. *R* genes confer direct fitness benefits via resistance and tolerance to pathogens [24–26]. However, there is a cost of resistance for growth [26,27], and positive and negative epistatic effects on plant fitness have been observed between *R* gene loci. For example, *RPS4* and *RRS1* function together and are genetically linked [28–31], and a small number of *R* gene combinations can trigger autoimmune hybrid necrotic interactions [32–34].

R genes belong to large paralogous gene families, that show high levels of diversity within and between species [16,17,35,36]. For example, the *Arabidopsis thaliana* NBS-LRR gene family consists of ~190–200 members, of which 30 have characterized resistance functions (*R* genes) (S1 Table) [16,35]. In this study we consider 197 genes in the NBS-LRR gene family (S1 Table). NBS-LRR genes show evidence of high historical recombination rates, inferred from patterns of linkage disequilibrium (LD) between single nucleotide polymorphisms (SNPs) [37–40]. NBS-LRR genes also show the signatures of both directional and balancing selection [40–46], with the latter thought to reflect negative frequency-dependent interactions [47], heterozygote advantage (overdominance) [48], or selection in diffuse or fluctuating environments [26]. Physical clustering of NBS-LRR genes along chromosomes is common and extensive experimental evidence for recombination exists, including the signatures of non-allelic gene conversion and unequal crossover [16,49–59]. This includes demonstration that unequal crossing-over within the maize *Rp1* cluster can generate novel resistance specificities not present in either parent [54].

Relatively little is known about variation in fine-scale crossover frequency between *R* genes. We therefore sought to use experimental genetics to measure meiotic crossovers associated with the Arabidopsis NBS-LRR gene family. We observe strong crossover hotspots in a subset of NBS-LRR genes. However, we also observe significant NBS-LRR gene coldspots, which in some cases associate with structural heterozygosity. We mapped fine-scale crossover patterns within the *RESISTANCE TO ALBUGO CANDIDA1* (*RAC1*) *R* gene hotspot, which confers resistance against oomycete pathogens [60–62]. *RAC1* intragenic recombination was found to overlap exons encoding the TIR, NBS and LRR domains. Highest *RAC1* crossover frequency occurs adjacent to the promoter and overlaps hotspot-associated CTT-repeat DNA sequence motifs [37,63,64]. Using a crossover assay based on flow cytometry of fluorescent pollen, which allows >10,000s post-meiotic gametes to be scored per individual [65–67], we show significant modification of NBS-LRR cluster recombination rate by genetic background. Using Southern blots we present data on NBS-LRR gene structural variation between ecotypes and relate this to patterns of crossover frequency. Together these analyses reveal heterogeneity in *R* gene crossover frequency and show how natural variation modifies NBS-LRR cluster recombination.

Results

Mapping crossover frequency within the Arabidopsis NBS-LRR gene clusters

NBS-LRR genes frequently cluster in plant genomes [16]. For example, the majority of the Arabidopsis NBS-LRR domain gene family (114/197) are located in clusters on the south arms of chromosomes 1 and 5 (Fig 1 and S1 Table), which also correspond to regions of disrupted synteny between *A.thaliana* and *A.lyrata* chromosomes [16,36]. We constructed phylogenies using NBS domain amino acid sequences and compared the resulting trees to the physical location of NBS-LRR genes on chromosomes 1 and 5 (Fig 1C and 1D). This indicated that both tandem and dispersed NBS-LRR gene duplications contribute to cluster organization, consistent with previous analyses [16,36]. To map crossovers experimentally at high resolution within the clusters we designed a double-selection strategy (Fig 1A). This utilizes Syngenta

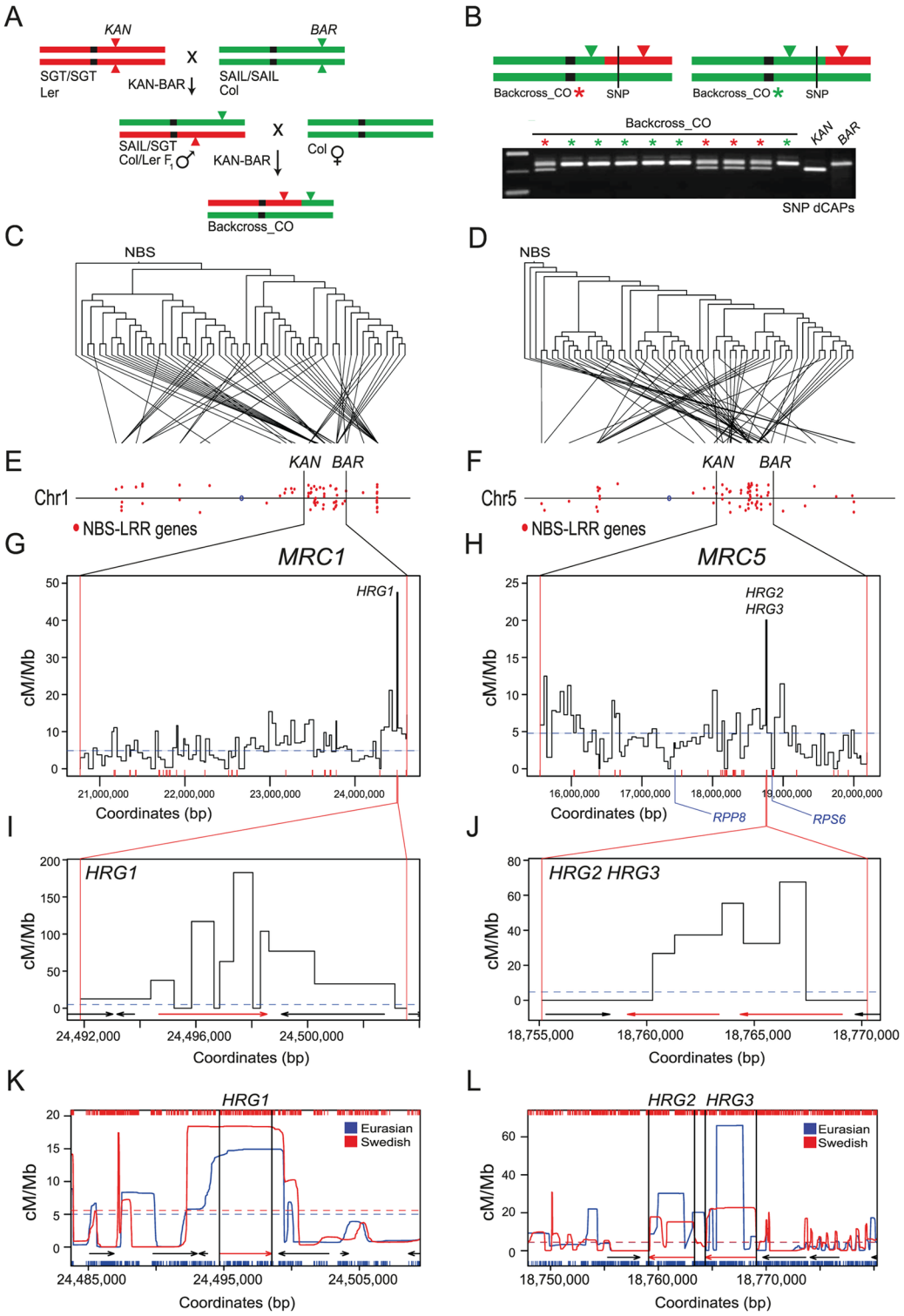


Fig 1. Meiotic recombination hotspots and coldspots within *MAJOR RESISTANCE CLUSTER1* and *MAJOR RESISTANCE CLUSTER5*. (A) Crossing scheme used to double-select *MRC* crossovers. *KAN* (SGT) Ler chromosomes are shown in red and *BAR* (SAIL) Col chromosomes are shown in green. Triangles represent insertions and are labeled with the resistance they confer (*KAN* = kanamycin, *BAR* = glufosinate). **(B)** Diagram illustrating *MRC* genotype expectations in double-resistant backcross progeny, depending on crossover location. A representative ethidium-stained genotyping gel for a Col/Ler single nucleotide polymorphism (SNP)

dCAPs marker is shown beneath [70]. Colored stars above the genotyped samples match the chromosome diagrams above. **(C-D)** Phylogenetic trees generated using NBS domain amino acid sequences are plotted, with branch tips connected to the corresponding gene physical location. **(E-F)** Physical maps of chromosomes 1 and 5 with red dots indicating NBS-LRR gene positions (S1 Table). The selected *MRC* regions are indicated by black vertical lines and *KAN* and *BAR* symbols. **(G-H)** Recombination rate (cM/Mb) throughout the genotyped *MRC1* and *MRC5* regions (S2 and S3 Tables). Red x-axis ticks indicate NBS-LRR gene positions (S1 Table). The dotted horizontal line indicates the male Col/Ler chromosome average recombination rate [71]. The *RPP8* and *RPS6* coldspot locations are indicated beneath the *MRC5* map. **(I)** Crossover frequency within the *HRG1* hotspot interval (S4 Table). Arrows represent genes, with the *R* gene highlighted in red. **(J)** As for (I) but for *HRG2*/*HRG3* (S5 Table). **(K-L)** Historical recombination rates across *HRG1* and *HRG2*/*HRG3*. Crossover frequency (cM/Mb) estimates (LDhat) from 80 Eurasian and 180 Swedish accessions are shown in blue and red respectively [72–74]. X-axis ticks indicate SNP positions used for analysis. Horizontal dotted lines indicate chromosome mean historical recombination values. Arrows represent genes, with the NBS-LRR genes highlighted in red.

doi:10.1371/journal.pgen.1006179.g001

Arabidopsis Insertion Library (SAIL) T-DNA insertions that confer glufosinate resistance in the Columbia (Col) accession, hereafter named *BAR* (Basta Resistant) lines [68], and Singapore Gene Trap (SGT) *Ds* insertions that confer kanamycin resistance in the Landsberg *erecta* (Ler) accession, hereafter named *KAN* (Kanamycin Resistant) lines (Fig 1A and 1B) [68,69]. Single copy *KAN* and *BAR* insertions that flank the NBS-LRR clusters on chromosomes 1 and 5 were identified using Southern blotting and genetic analysis (Figs 1 and S1). *KAN-BAR* insertion pairs used for crossover double-selection define the *MAJOR RESISTANCE CLUSTER1* (*MRC1*, 3.87 Mb, SGT_5517 and SAIL_148_D06) and *MAJOR RESISTANCE CLUSTER5* (*MRC5*, 4.69 Mb, SAIL_309_G06 and SGT_5183) map regions (Fig 1A and S2 and S3 Tables).

Paired *MRC KAN-BAR* insertions were crossed to generate double kanamycin-glufosinate resistant F_1 *trans*-hemizygotes, which were then backcrossed as males to wild type Col females (Fig 1A). Backcross progeny were grown on kanamycin-glufosinate media and the ratio of double-resistant to sensitive individuals was used to measure genetic distances, according to the formula $cM = 100\times(2\times(N_{KB}/N_{total}))$ (Fig 1A and Table 1), where N_{KB} is the number of kanamycin-glufosinate double-resistant progeny and N_{total} is the total number of progeny screened. Genotyping of double-resistant individuals with Col/Ler Simple Sequence Length Polymorphisms (SSLP) markers was used to confirm that they had experienced a crossover within the selected *MRC* region (Fig 1B and S2 and S3 Tables). Average recombination rates within the *MRC* regions were comparable to the male Col/Ler genome average 4.82 cM/Mb (*MRC1* 5.89 cM/Mb, *MRC5* 4.18 cM/Mb) (Table 1 and S2 and S3 Tables) [71]. To analyze fine-scale recombination patterns within the *MRC* regions a subset of the double-selected *MRC1* ($n = 725$) and *MRC5* ($n = 714$) recombinant individuals were chosen for further analysis by internal genotyping (Fig 1 and S2 and S3 Tables). Double-resistant crossover progeny were genotyped for 93 and 90

Table 1. Double-selection of *MRC1* and *MRC5* crossovers. *KAN-BAR* F_1 *trans*-hemizygotes were backcrossed to Col wild type female flowers and progeny were double-selected for kanamycin and glufosinate resistance. The numbers of double-resistant versus sensitive backcross seedlings are listed. The physical distances (megabase, Mb) were calculated between the *KAN* and *BAR* insertion sites relative to the TAIR10 Col reference assembly [75]. Genetic distance (cM) is calculated as $100\times(2\times(N_{KB}/N_{total}))$, where N_{KB} is the number of Kan-Bar double-resistant backcross progeny and N_{total} is the total number of backcross progeny screened (Double-Resistant+Sensitive).

	MRC1	***MRC5***
Double-Resistant	1,028	947
Sensitive	7,988	8,716
Total	9,016	9,663
cM	22.80	19.60
Mb	3.87	4.69
cM/Mb	5.89	4.18

doi:10.1371/journal.pgen.1006179.t001

Col/Ler SNPs within the clusters to identify *MRC1* and *MRC5* marker intervals containing crossovers, with a mean inter-marker distance of 46.5 kb (Fig 1 and S2 and S3 Tables). Analysis of crossover patterns in the *MRC* maps revealed intervals with both significantly higher and lower crossovers than expected if random (*MRC1*, X^2 = 441.45 d.f. = 93 $P = <2.2\times10^{-16}$; *MRC5*, X^2 = 337.29 d.f. = 90 $P = <2.2\times10^{-16}$), including intervals containing *R* genes (Fig 1 and S2 and S3 Tables).

NBS-LRR gene hotspots and coldspots within *MRC1* and *MRC5*

The *MRC1* interval with the highest recombination contains a functionally uncharacterized TIR-NBS-LRR gene (At1g65850), which for the purposes of this study we designate *HOT R GENE1* (*HRG1*) (Fig 1G and S2 Table). *HRG1* shares high sequence identity with *RPP1* and *DM2*, which confer resistance to *Hylaperonospora parasitica* and cause hybrid necrosis respectively [32,76]. The *HRG1* map interval contained 18 crossovers, which we fine-mapped by genotyping using derived Cleaved Amplified Polymorphisms (dCAPs) markers (Fig 1I) [70]. We observed a concentration of crossovers (14 of 18) within *HRG1*, with a peak recombination rate of 182.87 cM/Mb (Fig 1I and S4 Table), compared to the male Col/Ler genome average of 4.82 cM/Mb [71]. Similarly, the interval in *MRC5* with highest crossover frequency contains a tandem pair of TIR-NBS-LRR genes (At5g46260 and At5g46270), which we designate *HOT R GENE2* (*HRG2*) and *HOT R GENE3* (*HRG3*) (Fig 1H and S3 Table). *HRG2* and *HRG3* are also functionally uncharacterized, but show high sequence similarity with *RPS6* and *RAC1*, which confer resistance to *Pseudomonas syringae* and *Albugo candida* respectively [60,77]. Fine-mapping crossovers using additional dCAPs markers again revealed crossovers concentrated within the *HRG2* and *HRG3* NBS-LRR genes (11 of 11 crossovers) (Fig 1J and S5 Table). We fine-mapped crossovers in six additional high recombination intervals in *MRC1* and *MRC5* that contained NBS-LRR genes (Fig 2). We observed enrichment of crossovers within NBS-LRR genes in all cases (*WRR4*, *CW9*, *HRG4*, *HRG5*, *HRG6*, *HRG7*, *HRG8* and *HRG9*) (Fig 2 and S6–S11 Tables). Recombination rates observed at the *MRC* NBS-LRR hotspots were comparable to characterized plant hotspots and were detected at singleton, tandemly duplicated and complex NBS-LRR loci (Fig 2 and S4–S12 Tables).

To test for *MRC* intervals with significantly different crossover rates, we constructed 2×2 contingency tables and performed X^2 tests using the observed crossover and non-recombinant counts, and those expected at random based on the physical size of each interval. Finally, we applied a false discovery correction, due to multiple testing [78]. Out of 185 intervals in *MRC1* and *MRC5*, 22 were significantly different from the null expectation (S2 and S3 Tables). Ten intervals showed higher recombination, which included *HRG1*, *HRG2*, *HRG3*, *HRG4* and *HRG5* (S2 and S3 Tables). Twelve intervals showed lower recombination, which included intervals containing *RRS1B-RPS4B*, the *RPS6* cluster and a cluster of 4 uncharacterized CC-NBS-LRR genes (At1g58807, At1g58848, At1g59124 and At1g59218) (S2 and S3 Tables). However, after correction for multiple testing, only *HRG1* and the CC-NBS-LRR cluster remained significantly different (S2 and S3 Tables) [78]. To test whether intervals containing NBS-LRR genes showed different recombination rates compared to those that did not, we compared observed crossover counts to those expected at random in proportion to physical distance, using a X^2 test. *MRC* intervals that contained NBS-LRR genes did not have significantly different crossover counts compared to those that lacked NBS-LRR genes (P = 0.418).

As a further test of recombination in NBS-LRR genes compared to other genes, we analyzed 1,230 crossovers mapped throughout the genome by genotyping-by-sequencing of 192 Col×Ler F_2 individuals (S13 Table) [79,80]. The crossovers were mapped to a mean resolution of 947 bp, representing 1,164,432 bp in total, which significantly overlapped with genes (Block-

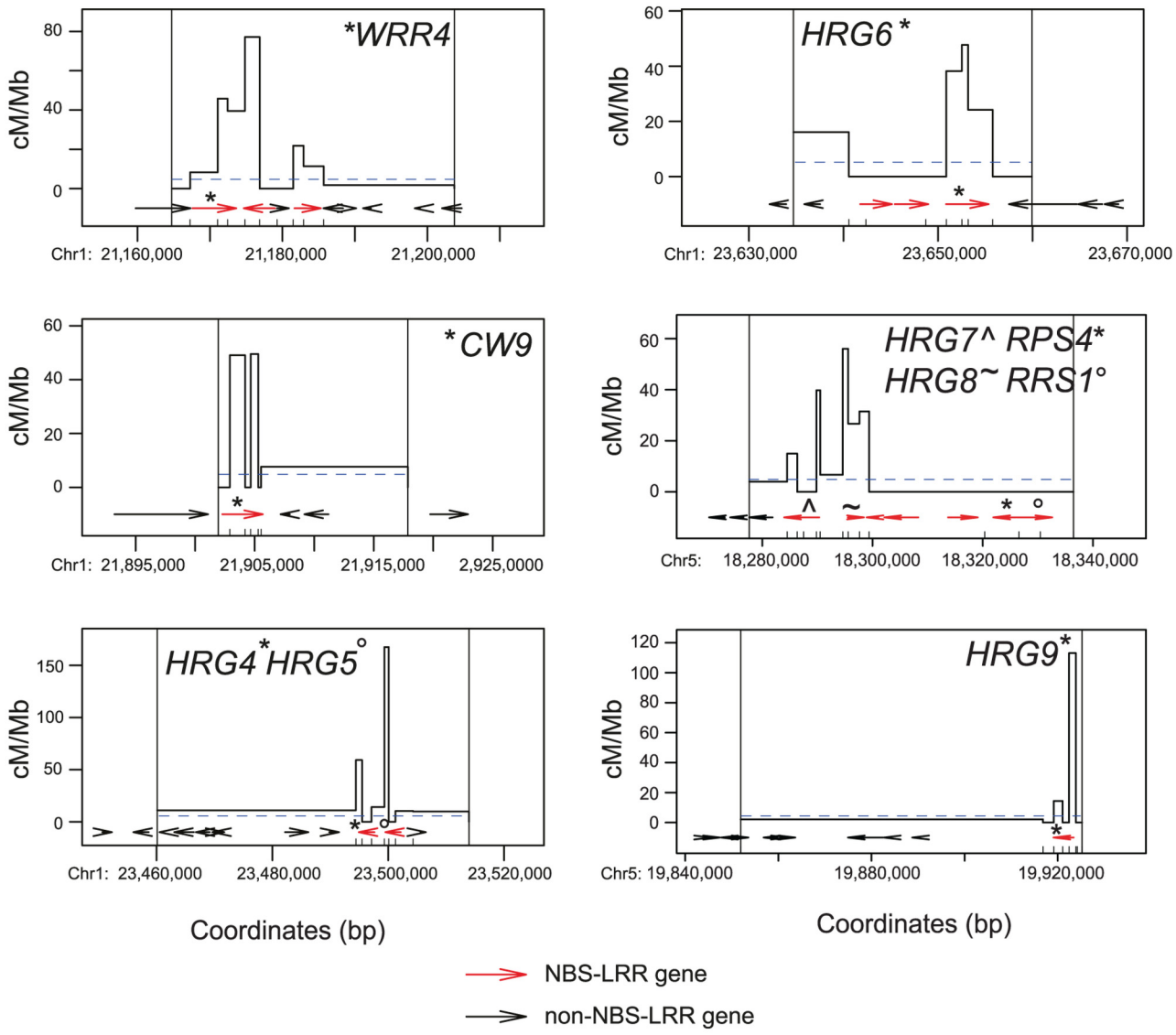


Fig 2. Fine-mapping *MRC* NBS-LRR gene crossover hotspots. Crossover frequency (cM/Mb) in high recombination *MRC* intervals containing NBS-LRR genes (S6–S11 Tables). Internal dCAPs marker positions are indicated by x-axis ticks, in addition to whether the region is located in *MRC1* (Chr1) or *MRC5* (Chr5). NBS-LRR genes are indicated by red arrows and non-NBS-LRR genes by black arrows. The chromosome average recombination rate for Col×Ler male backcrosses is indicated by horizontal, dotted blue lines [71]. Vertical black lines indicate the location of SNP markers genotyped in *MRC* maps (S2 and S3 Tables). NBS-LRR domain gene names are printed within the plot and symbols used to indicate which gene arrows they correspond to.

doi:10.1371/journal.pgen.1006179.g002

bootstrap Z = 4.54, $P = 2.77\times10^{-6}$). To compare gene groups we sampled from the non-NBS-LRR genes using parameters that approximated the clustering structure of the NBS-LRR genes (S2A and S2B Fig). We repeatedly sampled and tested for overlap between samplings and crossovers, requiring at least one base pair in common to score an overlap. Values were also length-normalized by rescaling the observed data by the total length in base pairs of the sample or gene set (S3C–S3F Fig). For both NBS-LRR and non-NBS-LRR genes the observed proportion of genes overlapping a crossover lie within those observed for the sampled data, usually near the distribution's center (S3C–S3F Fig). Therefore we conclude that the number of

NBS-LRR and non-NBS-LRR genes overlapping crossovers is not substantially different, and that both groups show the presence of hotspots and coldspots.

If recombination hotspots within NBS-LRR genes have persisted in the species, then decay of linkage disequilibrium (LD) between natural genetic polymorphisms is expected [81,82]. For example, the presence of all four possible gametes for pairs of linked polymorphisms (i.e. AB, aB, Ab and ab) can be used as a measure of historical recombination [81,82]. We therefore analyzed single nucleotide polymorphisms (SNPs) from 80 Eurasian and 180 Swedish Arabidopsis accessions using LDhat [72–74]. Consistent with our experimental crossover measurements *HRG1*, *HRG2* and *HRG3* show elevated historical recombination (Fig 1K and 1L) [72–74]. We also observed intervals containing *R* genes, like *RPP8*, with low experimental recombination, but which exhibited high historical recombination frequency (Fig 1H and S1–S3 Tables). The suppressive effect of structural polymorphism on recombination is likely to contribute to these patterns. For example, *RPP8* is tandemly duplicated in Ler relative to Col [50], and the *RPS6* cluster shows complex Col/Ler structural polymorphism [77]. In summary, although overall *MRC* recombination rates were close to the genome average, we observed heterogeneity in local crossover frequency and the presence of NBS-LRR gene associated hotspots and coldspots.

Intragenic recombination hotspots in the *RAC1* resistance gene

To investigate crossover variation within a single *R* gene we used pollen-typing [37,83,84], which is analogous to mammalian sperm-typing approaches [85,86]. This technique uses two rounds of nested, allele-specific PCR to specifically amplify and titrate either crossover or parental (non-recombinant) molecules from F_1 (e.g. Col/Ler) pollen DNA (Fig 3A and S3 Fig). Importantly, crossover products are not amplified from Col/Ler F_1 somatic leaf DNA under the same PCR conditions (Fig 3A). This demonstrates that template switching is not generating crossover molecules during PCR amplification. The concentrations of parental and crossover molecules, measured by titration, are used to calculate recombination rate within the amplified region [37,83–85]. Amplification from single molecules containing crossovers, followed by Sanger sequencing, is used to identify internal recombination locations, to the resolution of individual polymorphisms (Fig 3A) [37,83,84]. Between 100–200 crossover molecules are Sanger sequenced to map the topology of recombination within an amplified region and calculate crossover frequency (cM/Mb) for each marker interval (Fig 3A) [37,80,83,84]. Physical distances analyzed by pollen-typing are typically between 5–10 kb [37,80,83,84].

To select candidate hotspot *R* genes for pollen-typing analysis we analyzed historical recombination rates [37,72–74]. We used LDhat to analyse SNPs from 80 Eurasian and 180 Swedish Arabidopsis accessions [72,73], and observed that *RPP8* and *RAC1* are the *R* genes with highest LD-based recombination estimates (S1 Table). However, the *RPP8* locus shows structural polymorphism between accessions, including between Col and Ler, making it unsuitable for pollen-typing analysis [50]. In contrast, *RAC1* is present as a TIR-NBS-LRR singleton in both Col and Ler and so was selected for further study. *RAC1* alleles confer resistance to *Hyaloperonospora arabidopsidis* and *Albugo candida* and shows high sequence identity with *HRG2*, *HRG3* and *RPS6* [60–62]. *RAC1* resistance has been mapped using the Acem1 Albugo isolate, which both Col and Ler are susceptible to [60–62]. We amplified a 9.45 kb region, spanning *RAC1* and an adjacent gene (At1g31550) that encodes a putative lipase, and measured a recombination rate of 16.99 cM/Mb, which is comparable to known plant crossover hotspots (Fig 3C and S12 and S14 Tables). After sequencing 181 single molecules containing crossovers we observed highest recombination within the *RAC1* open reading frame, with a peak intragenic crossover frequency of 117 cM/Mb (Fig 3C and S14 and S15 Tables). *RAC1* recombination was highest in

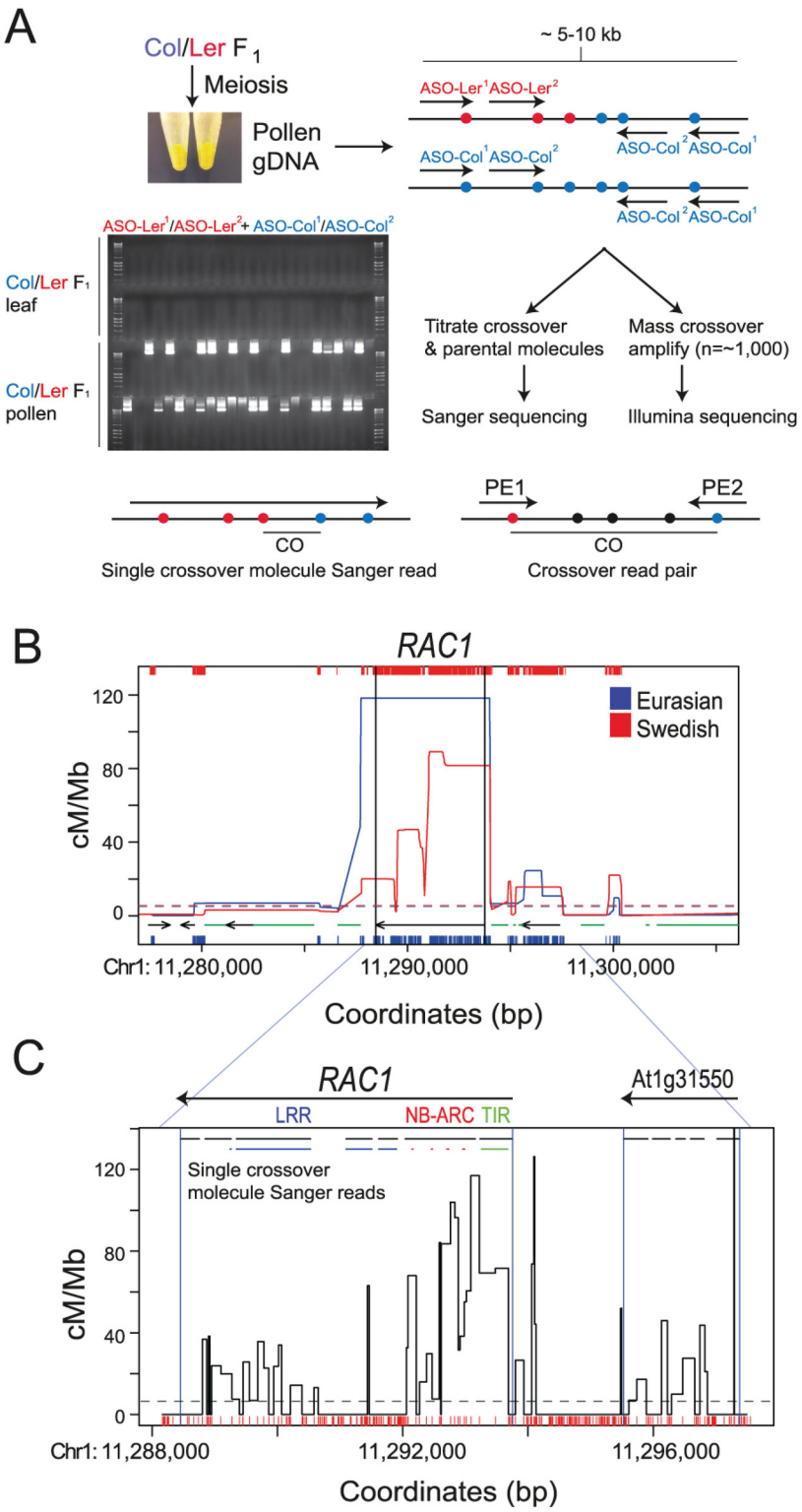


Fig 3. Intragenic crossover hotspots within the *RAC1* resistance gene. (A) Diagram illustrating pollen-typing methods used to fine-map plant crossover hotspots. Pollen is collected from Col/Ler F_1 plants and genomic DNA is used for two rounds of nested allele-specific PCR amplifications with either crossover (ASO-Ler+ASO-Col) or parental (ASO-Col+ASO-Col) allele-specific oligonucleotide (ASO) primer combinations. An ethidium-bromide stained agarose gel is shown with crossover-specific amplifications from

Col/Ler F_1 pollen, which are not observed when using Col/Ler F_1 leaf (somatic) DNA as a template. Crossover positions are identified using either Sanger sequencing from single crossover molecule amplifications, or paired-end sequencing of crossover molecule libraries. **(B)** Historical crossover frequency over *RAC1* estimated by LDhat analysis of polymorphisms from 80 Eurasian (blue) and 180 Swedish (red) accessions [72,73]. X-axis ticks indicate SNP positions used for analysis. Horizontal dotted lines indicate chromosome mean values. Black arrows represent genes and green horizontal lines indicate transposable elements. Black vertical lines indicate *RAC1* TSS and TTS positions. **(C)** Crossover frequency within the *RAC1* amplicon measured by crossover and parental molecule titration, followed by single crossover molecule Sanger sequencing. The positions of *RAC1* and the adjacent gene (At1g31550) are indicated by horizontal black arrows above the plot. Exons are indicated by horizontal black lines, with sequences encoding RAC1 TIR (green), NB-ARC (red) and LRR (blue) domains highlighted as coloured lines. For the NB-ARC domain, the position of the Walker A, Walker B, ARC1 and ARC2 motifs are shown. Red x-axis ticks show Col/Ler SNP positions. Blue vertical lines indicate the positions of gene TSS and TTS from TAIR10 representative gene models.

doi:10.1371/journal.pgen.1006179.g003

the 5' exons that encode the TIR and NBS domains, and in a 3'-region encoding LRR domains (Fig 3C). This pattern is consistent with high intragenic recombination rates observed at the 5' and 3' ends of plant genes [37,63,64,87]. In this regard it is important to note that residues critical for resistance function have been mapped to all three domains (TIR, NBS and LRR) in other *R* genes [58,88–91]. The open reading frame of the adjacent gene At1g31550 also showed substantial crossovers (Fig 3C). Together these data demonstrate that *RAC1* contains strong historical and contemporary crossover hotspots.

To analyze greater number of recombination events, we mass-amplified approximately 1,000 *RAC1* crossover molecules using 672 independent PCR reactions (Fig 3A). Each reaction was estimated to contain between 1 and 2 crossover templates, based on our previous titration experiments (S14 Table). The resulting crossover amplification products were then pooled, sonicated and used to generate sequencing libraries (Fig 3A). Paired-end sequencing was performed and read pairs were separated and aligned to the Col and Ler parental *RAC1* templates, allowing only exact matches (S16 Table). Reads were filtered for those that aligned to one parent only, and for pairs with a centromere-proximal read mapping to Col and a centromere-distal read mapping to Ler, on opposite strands (S16 Table). Read pairs matching these criteria are consistent with the crossover allele-specific oligonucleotide (ASO) configuration used during PCR amplification (Fig 3A). These filters yielded 182,909 total crossover read pairs (S16 Table). We assigned recombination values to marker intervals between crossover read pairs, weighted by the number of intervals and their widths, and normalized by the total number of mapping reads (Fig 4A). This revealed a similar recombination pattern to that observed from single molecule Sanger sequencing, with high recombination overlapping the *RAC1* TIR, NB-ARC and LRR domain exons (Figs 3C and 4A). Crossover read pair analysis also confirmed substantial recombination activity within the adjacent gene At1g31550 (Fig 4A). Together these data demonstrate focused intragenic recombination in *RAC1* overlapping exon sequences [60–62].

We investigated the extent to which chromatin and DNA sequences associate with crossover frequency at the *RAC1* hotspot. Previous studies have identified degenerate CTT and CCN repeat motifs associated with Arabidopsis hotspots, enriched at the 5′-end of genes [37,63,64]. We observed a CTT-motif (CTTCGTCATCTTCTTCT) almost immediately downstream of the *RAC1* start codon, within the TIR-domain encoding exon that shows high recombination rate (Fig 4A and 4B). CTT-motifs are also observed in *RAC1*-related NBS-LRR genes (Fig 4B). To investigate association of CTT-motifs with recombination we divided NBS-LRR genes located within the *MRC* maps into those in intervals with higher or lower crossover frequency, compared to the male Col×Ler genome average (4.82 cM/Mb) [71]. 62% (21 of 34) of high recombination NBS-LRR genes showed at least one match to a 16-mer CTT-repeat motif

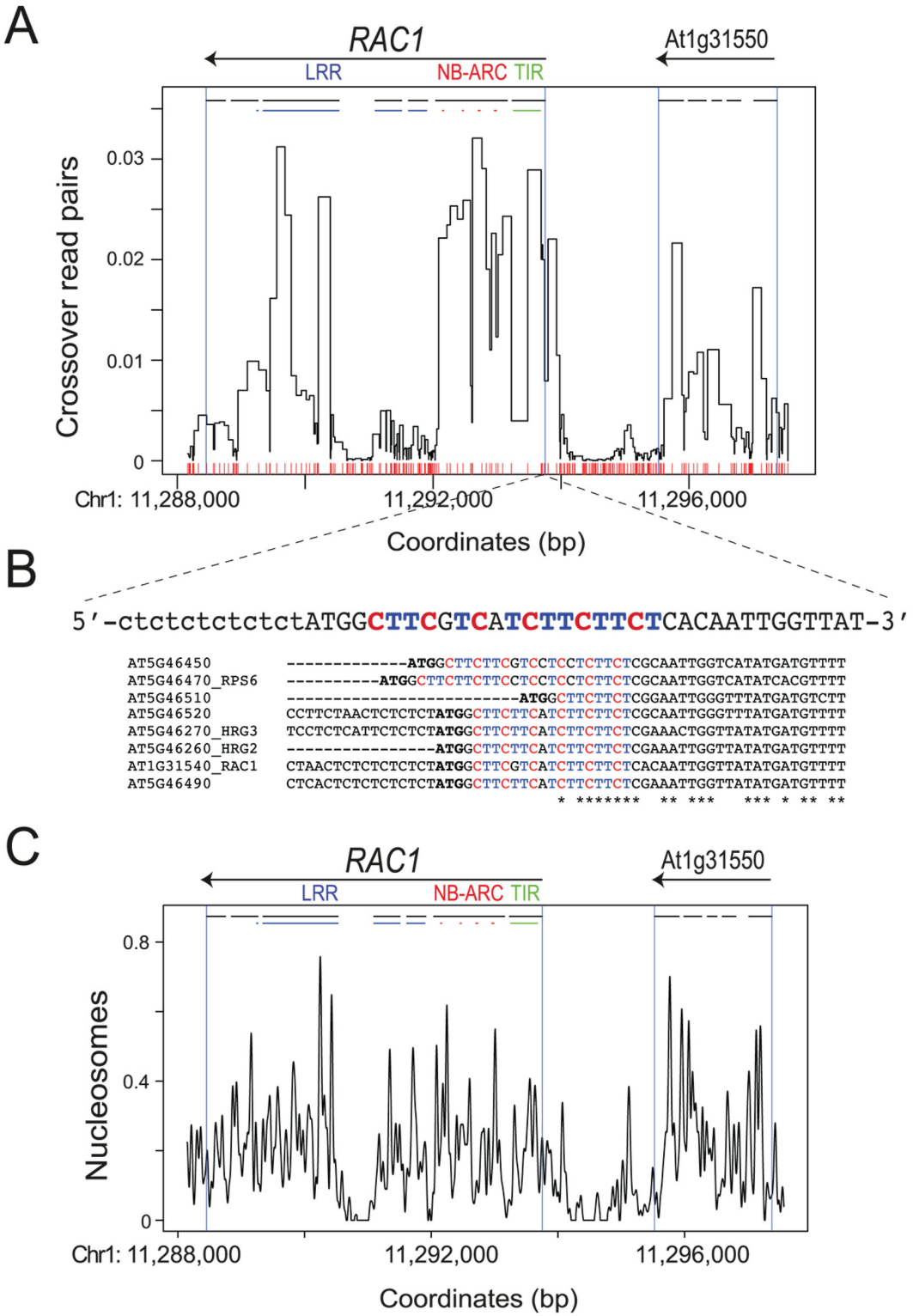


Fig 4. *RAC1* crossovers measured by high-throughput sequencing, nucleosome occupancy and CTT-repeat motifs. (A) Normalized crossover read pair values are plotted relative to the TAIR10 reference sequence, with the position of Col/Ler polymorphisms indicated by red x-axis ticks. Labelling is as for Fig 3C. **(B)** The sequence of the *RAC1* CTT-repeat motif is printed and connected to its physical location in (A) with dashed lines. Note the sequence is the reverse complement relative to the orientation of plotting in (A). The presence of CTT-

motifs in *RAC1*-related paralogs in the Col-0 accession are shown in a multiple sequence alignment, where gene start codons (ATG) are in bold type. The asterisks beneath the alignment indicate identical positions. **(C)** Normalized nucleosome occupancy values (MNase-seq) plotted as for (A).

doi:10.1371/journal.pgen.1006179.g004

that we previously identified as enriched at historical crossover hotspots (S17 Table) [37]. In contrast, significantly fewer (33%, 16 of 49) low recombination NBS-LRR genes showed matches to this CTT-motif (S17 Table) (X^2 $P = 0.016$). This analysis further demonstrates association between high crossover recombination and CTT-repeat motifs in Arabidopsis [37,39,63,64]. To analyse chromatin structure we performed micrococcal nuclease digestion of wild type (Col-0) chromatin and gel purified the resulting ~150 bp mononucleosomal DNA band [92]. This DNA was used to generate a library and paired-end sequencing performed, followed by analysis using the nucleR package [93]. We observed that intergenic regions within the pollen typing amplicon, and *RAC1* intron 4, showed relatively low nucleosome occupancy (Fig 4C) [92]. Highest crossover rates correlated with nucleosome-occupied exon sequences within *RAC1* and At1g31550 (Fig 4A and 4C). Together these analyses demonstrate associations between *RAC1* crossover frequency, DNA sequence motifs and chromatin structure.

R gene structural variation and crossover modification by genetic background

The *HRG1*, *HRG2-HRG3* and *RAC1* NBS-LRR hotspots show historical and experimental recombination rates above the genome average (Figs 1, 3, S4, S5, S14 and S15 Tables), whereas other *R* genes differed between these two measures. For example, *RPP8* shows high historical recombination, yet low experimental crossover frequency (Fig 1H and S3 Table). *RPP8* is a singleton *R* gene in the Col reference genome, but in Ler occurs as a tandemly duplicated pair with *RPH8A* [50]. Sequence analysis indicates that the Col *RPP8* haplotype is derived from a gene fusion event between *RPP8* and *RPH8A* [50]. We performed Southern blotting and hybridization analysis and confirmed the *RPP8/RPH8A* Col/Ler structural polymorphism (Fig 5A) [50]. The observed band sizes are also consistent with a recent Ler genome assembly generated by PacificBiosystems (Fig 5B) [94]. Hence, low *RPP8* recombination in the Col/Ler *MRC5* map may be explained due to structural heterozygosity inhibiting crossovers [95,96]. Similarly, the *RPS6* cluster shows structural polymorphism between Col and Ler and an absence of crossovers in *MRC5*, despite showing high sequence similarity to the hotspot *R* genes *RAC1*, *HRG2* and *HRG3* (Figs 1H, 5C and 5D and S3 Table) [77]. In contrast the hotspot *R* genes *HRG1*, *HRG2*, *HRG3* and *RAC1* show low levels of structural polymorphism between accessions, including between Col and Ler (Figs 5C and 5D and S4). For the *R* genes *RPP1* and *RPP4* Southern blotting and hybridization analysis revealed complex patterns of structural polymorphism that did not obviously correspond to the PacBio Ler assembly (S5 and S6 Figs) [94]. Hence, for these *R* genes structural polymorphism is also likely to be inhibitory to recombination, when divergent haplotypes are heterozygous.

To investigate the extent of inter-cross variation in NBS-LRR cluster recombination we used segregation of linked Fluorescent Tagged Line (FTL) T-DNA insertions [67,97]. Inheritance of FTLs in pollen can be monitored via fluorescence microscopy or flow cytometry (Fig 6A) [67,97]. Pollen grains collected from FTL/++ hemizygotes consists of four fluorescent classes that represent parental (no color and red+yellow) versus crossover recombinant (red or yellow alone) chromosomes (Fig 6B). The ratios between these fluorescent pollen classes are used to calculate genetic distance between the FTL T-DNA insertions (Fig 6A and 6B) [66,67,97]. Pollen fluorescence was analyzed using a flow cytometer allowing >10,000s of pollen grains to be scored per individual plant (Fig 6C and S18 Table) [67]. We used FTL lines that define a

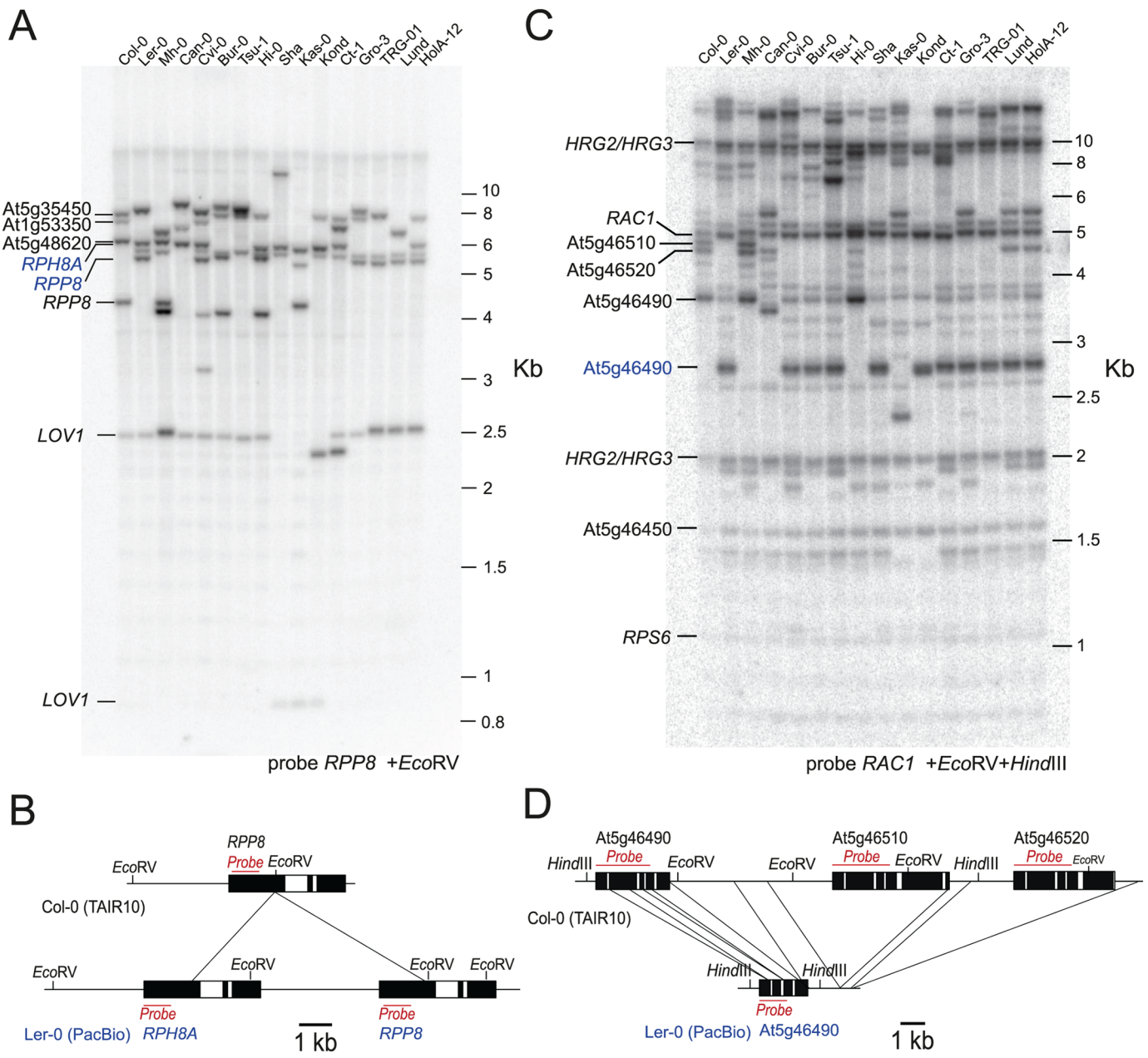


Fig 5. NBS-LRR gene structural diversity between Arabidopsis accessions. (A) Southern blotting and hybridization analysis of *RPP8* in the indicated accessions. Cross-hybridizing NBS-LRR genes in Col-0 predicted from the TAIR10 reference assembly are labelled in black. In Col-0 *RPP8* is detected as a 4.3 kb *Eco*RV band, and two 5.5 and 6.1 kb bands in Ler-0. Bands corresponding to Ler-0 *RPP8* and *RPPH8A* genes are labelled in blue. **(B)** Diagram showing annotated structural variation at *RPP8* between Col-0 (TAIR10) and Ler-0 (PacBio) assemblies [94]. Regions predicted to cross-hybridize with the probe are indicated in red, in addition to the position of *Eco*RV sites. **(C)** and **(D)** are as for (A) and (B), but probing for *RAC1*. The *RAC1* probe cross-hybridizes with a linked cluster of NBS-LRR genes, including *RPS6*. In Col-0 the *RPS6* cluster includes three NBS-LRR genes (At5g46490, At5g46510, At5g46520) that are present as a single gene (At5g46490) in Ler-0; such that in Col-0 three bands of 3.7, 4.5 and 4.7 kb are detected after *Eco*RV and *Hind*III digestion, and one 2.8 kb band in Ler-0. The Ler-0 band is labelled in blue.

doi:10.1371/journal.pgen.1006179.g005

4.92 Mb interval called *I5a*, which has a 2.09 Mb overlap with the *MRC5* region and contains 33 out of 70 chromosome 5 NBS-LRR genes (Fig 6A and S1 Table). We crossed *I5a* (generated in the Col-0 background) to 18 Arabidopsis accessions representing global diversity and measured genetic distance in replicate F_1 plants, scoring ~30,000 pollen grains per individual (Fig 6

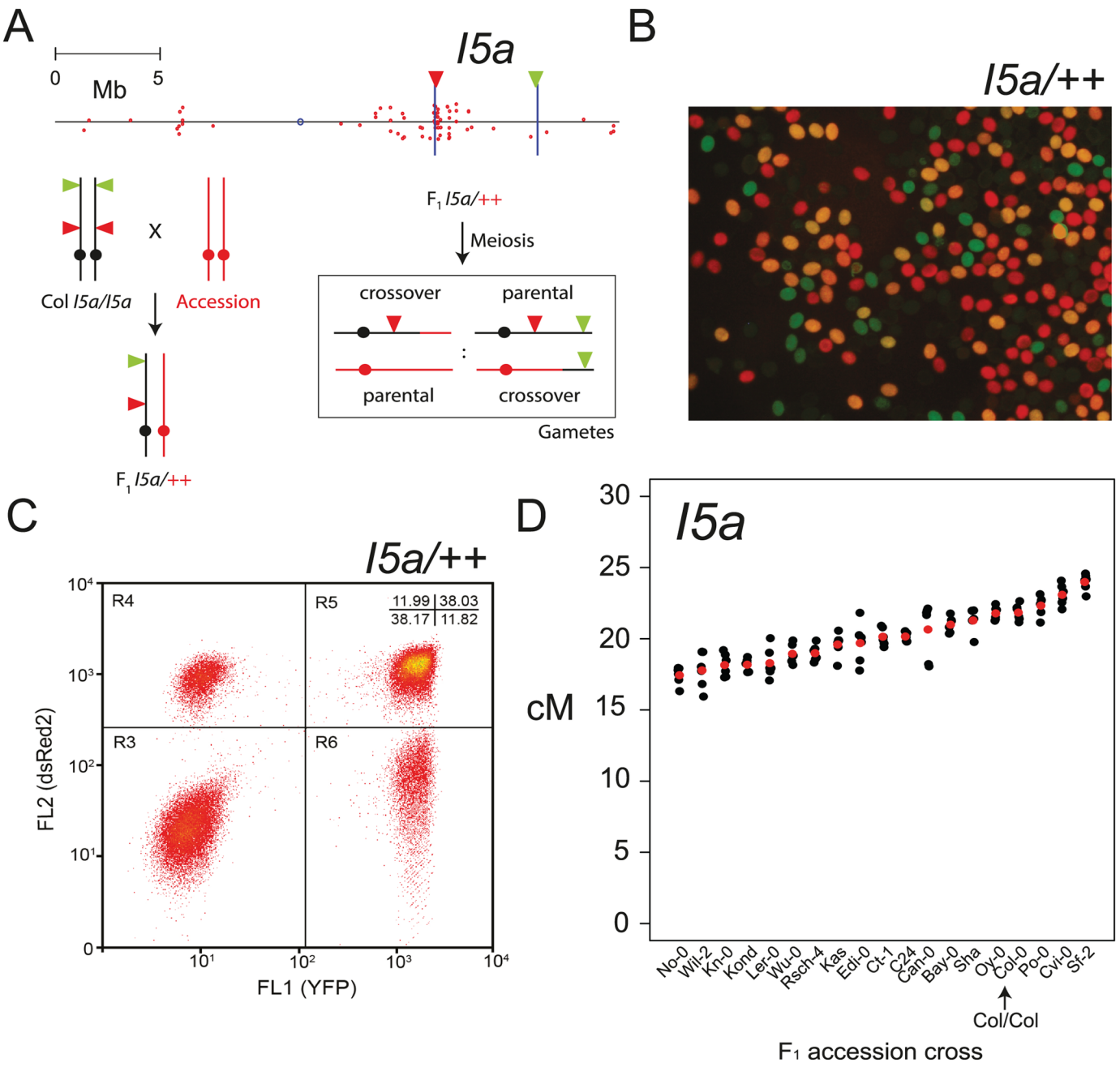


Fig 6. NBS-LRR cluster recombination rate is modified by natural genetic variation. (A) Physical map of chromosome 5 with NBS-LRR gene positions indicated by red dots. The positions of FTL T-DNAs that define *I5a* are indicated by vertical blue lines, with triangles indicating the colour (red or yellow) of fluorescent protein encoded. Beneath is a diagram illustrating the generation of FTL hemizygous F_1 plants and scoring of crossover recombination in pollen. Crossover within *I5a* leads to pollen with red or yellow fluorescence alone, whereas the parental classes are either without fluorescence or show both red and yellow. **(B)** Fluorescence micrograph of pollen from *I5a*/++ hemizygotes, showing the presence of four fluorescent classes. **(C)** A representative flow cytometry plot of pollen from *I5a*/++ hemizygotes. Points represent individual pollen grains measured for yellow (FL1 YFP) and red (FL2 dsRed2) fluorescence. **(D)** *I5a* genetic distance (cM) measured from individual F_1 plants (S18 Table). Mean values are shown in red. The accession crossed to is indicated on the x-axis. The Col/Col homozygous sample is labelled and indicated with an arrow.

doi:10.1371/journal.pgen.1006179.g006

and S18 Table). We observed 3 F_1 crosses with significantly higher *I5a* recombination and 15 F_1 crosses with significantly lower *I5a* recombination, relative to Col/Col homozygotes (general linearized model $P<2.0\times10^{-16}$) (Fig 6D and S18 Table). This demonstrates the sensitivity of NBS-LRR cluster recombination rates to genetic background, which is consistent with observed *cis* and *trans* modification of crossover frequency by natural genetic variation in plants [97–105]. Therefore, Arabidopsis genetic variation has a significant modifying effect on crossover frequency within the NBS-LRR clusters.

Discussion

The 'two-speed' genome hypothesis predicts recombination hotspots associated with disease resistance genes, due to the beneficial effects of diversity during pathogen coevolution [8–11]. Consistent with this we observed a subset of strong *R* gene hotspots. For example, *RAC1* shows high historical and contemporary crossover frequency, comparable to known plant hotspots (S12 Table). However, not all NBS-LRR genes are hotspots and a substantial number of coldspots were observed. Coldspot *R* genes in the *MRC* maps correlated with structural heterozygosity in some cases, for example, *RPS6* and *RPP8*. Heterogeneity between NBS-LRR genes is well documented in (i) patterns of sequence diversity and evolutionary history [44,107], (ii) the degree of sequence exchange observed between paralogs [57,59,106,107], (iii) direct versus indirect modes of effector recognition [13,20] and (iv) the extent of gene clustering [108]. Here we additionally demonstrate heterogeneity in Arabidopsis NBS-LRR gene meiotic recombination rates. The ecology, generation time and effector complement of a given pathogen will interact to determine the selection pressures on NBS-LRR genes. It is also important to consider that domains required for recognition versus downstream resistance signaling within a single NBS-LRR gene may be under distinct forms of selection, e.g. diversifying versus purifying. Hence, for a given *R-Avr* interaction, higher or lower recombination rates, at varying scales, may be beneficial. Relevant to these considerations is the degree of outcrossing in Arabidopsis, which is estimated between 1–3% [39,109]. For example, genotyping of local stands of Arabidopsis in Germany detected clear evidence for recent outcrossing and heterozygosity [110]. The rapid decay of LD in Arabidopsis also indicates significant levels of historical outcrossing [39,72,73,111].

We observed pronounced intragenic variation in *RAC1* crossover frequency, with TIR, NB-ARC and LRR encoding exons showing high recombination. It is important to note that all three domains have been shown to control resistance specificity and sensitivity in other *R* genes [88–91]. We propose that intragenic recombination patterns can be selected through a combination of sequence and epigenetic determinants. For example, CTT and CCN repeat sequence motifs associate with high recombination in Arabidopsis [37,63,64]. Similarly, we observed that CTT-motifs were enriched in high recombination NBS-LRR genes located in the *MRC* maps. These motifs co-localize with the +1 H2A.Z histone variant-containing nucleosome and also significantly overlap with historical and experimental recombination [37,63,64]. CTT/CCN motifs bear similarity to C-rich motifs bound by mammalian PRDM9 proteins, which determine hotspot locations via histone $H3K4^{me3}$ [112–117]. However, no clear homolog of PRDM9 exists in plants, and so CTT/CCN repeat motifs likely function via a different mechanism [7]. Association of high recombination and AT-rich sequences upstream of transcriptional start sites is also observed in Arabidopsis, which correlates with regions of low nucleosome occupancy [37,39,63,64]. Hence, Arabidopsis hotspot-associated DNA sequence motifs may function directly to recruit recombination-promoting factors, or alter chromatin and thereby influence accessibility of recombination factors. Extensive evidence exists for *cis* and *trans* acting genetic modifiers of meiotic recombination in plants, which also have the

potential to influence *R* gene crossovers [97–105]. The location of *R* genes along the chromosome is also likely to be a significant factor, as recombination rate shows broad scale variation, in addition to sex-specific patterns in the sub-telomeric regions [37,71,103,118].

Varying recombination rates within and between *R* genes has important consequences for the generation of recognition diversity, as gene conversion and unequal crossover have been associated with changes to *R* gene structure and function [16,49–59]. Heritable gene conversion events occur during meiosis and are associated with both crossover and non-crossover DNA double strand break repair [119,120]. Extensive NBS-LRR sequence analysis indicates gene conversion events within and between *R* paralogs, which are associated with *Avr* recognition differences. For example, in the tomato *Cf-4/9* and *Cf-2/5* clusters [49,121], the rice *Xa21* cluster [122], the lettuce *Dm3* cluster [51,123], the Arabidopsis *RPP8* and *RPP5* genes [50,52,53,106] and the Flax *M* and *N* loci [59,124]. Hence, higher recombination rates are likely to promote generation of functional allelic diversity via associated gene conversion events. Unequal crossovers contribute to changes in *R* gene resistance function by fusing, duplicating or deleting genes, or clusters of genes [49,51,53,54,56,59,125]. This is believed to occur due to misalignment of related, but non-allelic, sequences during strand invasion, followed by crossover recombination. Unequal recombination has been directly observed between the tomato *Cf4* and *Cf9* genes, which encode membrane bound LRR-repeat proteins that control resistance to *Cladosporium fulvum* [49,121,126]. Interestingly, increased levels of unequal recombination were detected in *Cf-4/Cf-9* heterozygotes compared with *Cf-9/Cf-9* homozygotes, suggesting that polymorphism can promote this type of recombination [49]. Similar testcrosses were performed with the maize *Rp1* cluster, which confers resistance against *Puccinia sorghi*, and significant levels of crossover-associated resistance change were observed using both heterozygotes (*Rp1-J1/Rp1-J3*) and homozygotes (*Rp1G/Rp1G*) [54–56,127]. Finally, the lettuce *Dm3* cluster that confers resistance against *Bremia lactucae* was found to show low recombination rates, though deletions associated with crossovers were observed [51].

Besides unequal recombination, canonical crossovers within *R* genes have the potential to generate new resistance functions, as amino acids controlling recognition specificity or signaling are known to occur in TIR, NBS and LRR protein domains [13,58,88–91,128]. As *R* genes tend to cluster, crossovers within these clusters can also generate or break beneficial resistance haplotypes, for example when linked *R* gene alleles function together [28–31]. High recombination rates are predicted to lead to high levels of structural diversity. However, we demonstrate that the *RAC1* resistance gene shows high recombination, yet occurs as a singleton TIR-NBS-LRR gene in all Arabidopsis ecotypes tested. In flax the *L* resistance gene is similar to *RAC1* in showing stability as a singleton TIR-NBS-LRR gene in all backgrounds, but with a high diversity of resistance alleles [57]. *RAC1* and *L* are both closely related to complex, tandemly duplicated loci that include *RPS6* and *M* respectively [57,59,60,77]. Hence, it will be interesting to further define how contrasting patterns of sequence diversity are maintained at related *R* gene loci and how this relates to meiotic recombination.

R gene clusters are reminiscent of mammalian major histocompatibility complex (*MHC*) regions [129]. For example, the 3–4 Mb *MHC* region on human chromosome 6 contains 240 genes, including multiple human leukocyte antigen (*HLA*) genes [129]. Both *MHC* and *R* cluster regions show high sequence diversity, the signatures of balancing selection and high non-synonymous:synonymous substitution ratios [38,40,42,44–46,129,130]. Although human genetic maps generally show low recombination in *MHC* regions, punctate hotspots have been observed (e.g. *DNA3*, 140 cM/Mb) [131]. Interestingly, virulence effectors have also been observed to cluster in the genomes of eukaryotic pathogens [132]. For example, the genome of the oomycete *Phytophora infestans* (potato blight) shows conserved gene-dense regions, interspersed with diverse repetitive blocks where virulence effectors are located [9,132], and the

basidiomycete fungus *Ustilago maydis* genome contains clustered genes encoding secreted virulence proteins that are transcriptionally induced during infection [11]. Similarly, the human pathogens *Trypanosoma brucei* and *Plasmodium falciparum* show sub-telomeric clustering of variant surface glycoproteins [133,134]. We speculate that coevolutionary dynamics drive mirrored changes in host and pathogen genomes and clustering of selected genes. It will be important to further explore inter-relationships between recombination, selection and plant-pathogen ecologies and how they influence diversity and gene organization at host resistance and pathogen effector loci.

Materials and Methods

NBS-LRR gene identification and analysis

We combined published annotations of the Arabidopsis NBS-LRR family with our own manual searches and selected a set of 197 genes for analysis (S1 Table) [16,35]. To analyse phylogenetic relationships NBS domain amino acid sequences were aligned using ClustalOmega and gaps treated as missing data. The resulting alignment was analysed using the PhyML 3.0 webserver (http://www.atgc-montpellier.fr/phyml/) with 500 bootstraps performed for tree construction [135]. To search for CTT-repeat motifs associated with NBS-LRR genes, a position weight matrix from a historical hotspot-enriched motif was matched at 80% to 1 kb windows surrounding NBS-LRR gene transcriptional start sites (TSSs).

Double-selection and analysis of *MRC1* and *MRC5* crossovers

To select for crossovers with the NBS-LRR gene clusters we followed an identical strategy to that used previously for the *TEL1* interval (Fig 1A) [37]. SAIL T-DNA insertions conferring glufosinate resistance in the Col accession (hereafter named *BAR*), and *Ds* insertions conferring kanamycin resistance in the Ler accession (hereafter named *KAN*), were chosen with insertion sites flanking the NBS-LRR gene clusters on the south arms of chromosomes 1 and 5 [68,69]. Southern blotting and hybridization was used to identify single copy *KAN* and *BAR* insertion lines (S1 Fig). Single copy lines were tested that (i) self-fertilization of homozygotes gave ~100% resistant progeny, (ii) self-fertilization of hemizygotes gave ~75% resistant progeny and (iii) backcrossing hemizygotes to wild type gave ~50% resistant progeny (S1 Fig). After screening we selected SGT_5517 (*KAN*) and SAIL_148_D06 (*BAR*) for chromosome 1 and SAIL_309_G06 (*BAR*) At5g49220 and SGT_5183 (*KAN*) for chromosome 5. *KAN* and *BAR* lines were then crossed to generate *trans*-hemizygote F_1 individuals, which were backcrossed as males to Col-0 (Table 1). Backcross progeny were germinated on media containing kanamycin and glufosinate to select for individuals that experienced a crossover between the *KAN* and *BAR* insertions. Genetic distance was calculated as: $cM = 100\times(2\times(N_{KG}/N_{total}))$ (Table 1), where N_{KG} is the number of Kan-Bar double-resistant backcross progeny and N_{total} is the total number of backcross progeny screened. We used Col/Ler SSLP genotyping markers to confirm that surviving progeny had experienced a crossover between the *MRC* insertions. Recombinants were selected using SSLP markers to restrict analysis to the *MRC1* and *MRC5* intervals (S2 and S3 Tables). Genomic DNA from these individuals was then genotyped using sets of Col/Ler SNP markers within *MRC1* and *MRC5* (KBiosciences). To test whether crossover distributions were significantly different from random we performed a $2\times n$ X^2 test, where n = the number of map intervals. To test the significance of individual intervals X^2 tests were performed between the observed crossover and non-recombinant counts with those expected at random using 2×2 contingency tables (S2 and S3 Tables). Random expectations were obtained by proportionally allocating crossovers to intervals according to their physical length, i.e. an

interval twice as long would be expected to receive twice as many crossovers. Finally, a false discovery correction was applied due to multiple testing [78].

Genotyping by sequencing and crossover analysis

DNA was extracted using CTAB and used to generate sequencing libraries as described [79,80], with the following modifications. DNA was extracted from 3 rosette leaves of 5 week old plants and 150 ng of DNA used as input for each library. DNA shearing was carried out for 20 minutes at 37°C with 0.4U of DNA Shearase (Zymo research). Each set of 96 libraries was sequenced on one lane of an Illumina NextSeq500 instrument (300-cycle Mid Output run). Sequencing data was analysed to identify crossovers as previously reported, using the TIGER pipeline [79,80]. The fastq files for these data have been uploaded to ArrayExpress accession E-MTAB-4657.

To compare gene groups we sampled from the non-NBS-LRR genes by taking n blocks of length b randomly distributed in the genome and sampled a proportion p of genes within each block. We compared the distributions of the distance between the start coordinates of the sampled genes and the physical widths of the sampled genes, to those observed in the NBS-LRR genes, performing parameter sweeps on n, b and p, in order to minimize the sum of the Kolmogorov-Smirnov statistics observed for each distribution. The minimal difference between the distributions was observed for 220 blocks of 35,000 base pairs and sampling 20% of the genes in each block. Using these parameters, we obtained an approximation to the clustering structure of the NBS-LRR genes (S2A and S2B Fig). We then repeatedly sampled (3,200 times) using these parameters and tested for overlap between samplings and crossovers, requiring at least one base in common to score an overlap. Values were length-normalized by rescaling the observed data by the total length in base pairs of the sample or gene set (S3C–S3F Fig). The observed overlap between NBS-LRR and non-NBS-LRR genes with crossovers were then compared to the sampled data (S3C–S3F Fig).

RAC1 hotspot pollen-typing and crossover sequencing

The *RAC1/RPP9* gene and surrounding region were PCR amplified and Sanger sequenced from the Ler-0 line used in our laboratory to identify all associated SNPs and indel polymorphisms. Using this information we designed allele specific oligonucleotides (ASOs) that specifically amplify from either Col or Ler genomic DNA templates, as described previously [37,83,84,136] (S1 Fig and S19 Table). For *RAC1* two nested ASO amplifications were performed with the following sets of ASO primer pairs to amplify either crossover (Ler-Col) or parental (Col-Col) molecules:

1st Crossover amplification = KC459 + KC418

1st Parental amplification = KC493 + KC418

2nd Crossover amplification = KC465 + KC417

2nd Parental amplification = KC495 + KC417

The same PCR conditions were used for 1st and 2nd PCRs, which were:
94°C for 2'30",
then 5 cycles of {94°C for 30", 65°C (decrease by 0.4°C per cycle) for 45", 68°C for 8'00"},
then 25 cycles of {94°C for 20", 63°C for 30", 68°C for 8'00"},
then {68°C for 10'00"}.

Recombination rate was estimated by titration of crossover and parental molecule concentrations from F_1 pollen genomic DNA [37,83,84,136]. DNA concentrations were identified which

gave approximately half crossover positive amplifications, where by Poisson approximation the majority are from single molecules [37,83,84,136]. Single crossover molecules were amplified and treated with exonuclease I (NEB, M0293) and shrimp alkaline phosphatase (Amersham, E70092), and then directly Sanger sequenced to identify internal crossover locations.

Based on crossover molecule concentrations measured by titration, we performed mass amplification (672 reactions) of crossovers, such that most contained between 1–2 crossover templates. The resulting amplification products were pooled, concentrated by isopropanol precipitation and gel purified. ~2 μg of DNA in 100 μl of TE (10 mM Tris, 1mM EDTA, pH 8.0) buffer was then sonicated to a size range of 300–400 bp using a Bioruptor (Diagenode). Sonicated DNA was gel purified, end-repaired and used to generate libraries (Tru-seq, Illumina). These libraries were sequenced on an Illumina nextSeq instrument using paired-end 75 bp reads. Paired reads were separated and aligned to the Col or Ler parental sequences using bowtie allowing only exact matches [137]. Reads were filtered for those that uniquely aligned to one parental sequence only. We then filtered for read pairs that had a centromere-proximal match to Col and a centromere-distal match to Ler on opposite strands, in order to identify crossover read pairs (S16 Table). These criteria are consistent with the ASO configuration used during allele-specific PCR amplification. A value of 1 was divided by the distance (bp) between each crossover read pair. For example, if a given pair of crossover reads are 200 bp apart, then the fractional value per bp is 1/200 = 0.005. Further consider 2 SNPs located between these crossover reads that divide this region into 3 intervals of 10, 50 and 140 bp. Each interval would then be assigned values of 10*0.005, 50*0.005 and 140*0.005 respectively. This process is repeated for all read pairs and final values normalized by the total number of crossover read pairs. The *RAC1* pollen-typing fastq files have been uploaded to ArrayExpression accession E-MTAB-4556.

Analysis of historical crossover frequency

Diallelic SNPs with less than 10% missing data were selected from 80 Eurasian and 180 Swedish accession 1,001 Genomes Project datasets [72,73]. SNPs overlapping repeats and the centromeres were masked due to difficulties in assembling these regions, as performed previously [37]. SNP data was analysed using LDhat run parameters described previously, but using a $\theta = 0.001$ look-up table [37,74]. LDhat recombination maps were filtered by setting regions where $4N_e r > 100$, or the distance between SNPs >50 kb, to zero recombination. Population-scaled recombination rates were rescaled to cM/Mb by regression onto a consensus experimental F_2 genetic map, as described previously [37].

Micrococcal nuclease sequencing

Chromatin from 1 gram of wild type (Col-0) flowers (unopened buds, prior to floral stage 12) was isolated and digested with 0.05 units of mirococcal nuclease (MNase, NEB) in reaction butter (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 1 mM EDTA, 4 mM $CaCl_2$) at 37°C for 15 minutes, with occasional vortexing. The digested nucleosomal DNA was then treated with proteinase K (20 μg/μl) and phenol/chroloform extraction performed, followed by ethanol precipitation. The digested DNA was separated using gel electrophoresis through an ethidium-stained agarose-TBE gel and the ~150 bp nucleosomal band was gel purified. This DNA was end-repaired and used to generate Illumina sequencing libraries (Tru-seq) and subjected to paired-end 100 bp sequencing on a Hi-seq instrument. Reads were mapped using Bowtie2 to the TAIR10 assembly and normalized by library size. Nucleosome occupancy was analysed using the nucleR package [93]. The MNase-seq fastq files have been uploaded to ArrayExpression accession E-MTAB-4556.

NBS-LRR gene Southern blotting and hybridization

1.5 μg of genomic DNA was digested using 100 units of restriction enzyme (NEB) in a final volume of 200 μl for 15 hours at 37°C. Digested DNA samples were isopropanol precipitated, resuspended in TE and electrophoresed in 0.8% agarose-TBE gels. Gels were blotted onto a positively charged nylon membrane (Hybond-N+, Amersham Biosciences), which was hybridized in 0.5 M phosphate buffer, 7% w/v SDS, 1 mM EDTA (pH 8) and 1% BSA at 65°C. DNA probes to the *RAC1* gene (a PCR fragment amplified with RAC1-F 5′-TGATAGAGATTGAGG CGGTCA-3′ and RAC1-R 5′-GGCTTCGTCATCTTCTTCTCAC-3′), *RPP8* (a PCR fragment amplified with RPP8-F 5′-TCTTCATTCCTTCAATCTTCAGT-3′ and RPP8-R 5′-GAGAT ACTTGAGTTTATACGAGGCTAA-3′), *HRG1* (a PCR fragment amplified with HRG1-F 5′-GGTTTGGTCATGGAAGTCGG-3′ and HRG1-R 5′-TCTTACCCGGTTTCCTTCCC-3′), *RPP1* (a PCR fragment amplified with RPP1-F 5′-AGTCAGCAAGAGGAGACCAC-3′ and RPP1-R 5′-TCCAACTTAGTGCAATCATGGTC-3′) or *RPP4* (a PCR fragment amplified with RPP4-F 5′-AGAAGGCAAACGCTTCACTG-3′ and RPP4-R 5′-AGCTGCCATCTCAAGGT CTT-3′) were labelled with [α-32P]-dCTP using a random priming labelling kit (Rediprime II DNA labelling system, Amersham). Blots were washed with 0.5% SSC, 0.1% SDS solution at 65°C and imaged with a PhosphoImager (Typhoon 8610, Molecular Dynamics). We compared Southern blot data with the TAIR10 Col assembly and a Ler assembly generated by Pacific Biosystems (http://blog.pacificbiosciences.com/2013/08/new-data-release-arabidopsis-assembly.html) [94].

Measurement of crossovers using flow cytometry of fluorescent pollen

Analysis of fluorescent pollen and segregation of *I5a* FTL T-DNAs was performed as described previously [67]. Genetic distance is calculated as cM = 100×(R5/(R3+R5)), where R5 is the number of yellow-alone pollen and R3 is the number of both yellow and red fluorescent pollen. To test whether recombinant and non-recombinant counts were significantly different between replicate groups we used a generalized linear model (GLM). We assumed the count data is binomially distributed:

$$Y_i \sim B(n_i, p_i)$$

where Y_i represents the recombinant counts, n_i are the total counts, and we wish to model the proportions Y_i/n_i. Then:

$$E(Y_i/n_i) = p_i$$

and,

$$var(Y_i/n_i) = \frac{p_{i(1-p_i)}}{n_i}$$

Thus, our variance function is:

$$V(\mu_i) = \mu_i(1-\mu_i)$$

and our link function must map from (0,1) -> (-∞, ∞). We used a logistic link function which is:

$$g(\mu_i) = \text{logit}(\mu_i) = \log\frac{\mu_i}{1-\mu_i} = \boldsymbol{\beta X} + \boldsymbol{\varepsilon}_i$$

where $\boldsymbol{\varepsilon}_i \sim N(0,\sigma^2)$. Both replicates and genotypes are treated as independent variables ($\boldsymbol{X}$) in our model. We considered P values less than 0.05 as significant.

Supporting Information

S1 Fig. Identifying single locus KAN and BAR lines for MRC genetic mapping. (A) Representative Southern blots used to identify single copy SAIL and SGT lines. Replicate samples indicate DNA isolated from siblings. Single copy lines were identified for further experiments. **(B)** Self-fertilization of homozygous KAN or BAR lines should yield ~100% resistance progeny on selective plates. **(C)** Self-fertilization of hemizygous KAN or BAR lines should yield 75% resistant progeny on selective plates. **(D)** Backcrossing hemizygous KAN or BAR lines should yield 50% resistant progeny on selective plates.
(TIFF)

S2 Fig. Overlap of crossovers with NBS-LRR genes compared with other genes. (A) The distribution of distances between start coordinates for NBS-LRR genes (black) and sampled data (blue). **(B)** As for (A), but showing the distribution of gene widths. **(C)** Boxplot showing the distribution of the number of crossovers per gene from the sampled data. The red point shows the observed value for NBS-LRR genes and the blue point shows that for the non-NBS-LRR genes. **(D)** As for (C), but showing the number of crossovers overlapping the sampled genes, normalized by the total length of all genes in the sample. **(E)** Boxplot showing the distribution of the proportion of genes overlapping crossovers in the sampled data. The red point shows the observed value for NBS-LRR genes and the blue point shows that for the non-NBS-LRR genes. **(F)** As for (E), but showing the number of genes overlapping crossovers normalized by the total length of all genes in the sample.
(TIFF)

S3 Fig. Validation of template specificity for *RAC1* allele-specific oligonucleotides used for pollen-typing. Representative ethidium bromide stained agarose gels are shown with PCR amplification products generated using either Col or Ler genomic DNA as a template. Reactions were repeated using the gradient of annealing temperatures indicated at the top of the gel. Primers were either template non-specific (universal forward (UF) or universal reverse (UR)) or allele-specific (Col R1 = KC418, Col R2 = KC417, Ler F1 = KC459, Ler F2 = KC465). Primer sequences are provided in S19 Table. Amplification was specific to the template genotype of choice under these PCR conditions.
(TIFF)

S4 Fig. *HRG1* structural diversity between Arabidopsis accessions analysed by Southern blotting and hybridization. Genomic DNA was isolated from the indicated accessions and digested with *Eco*RV. DNA was separated using gel electrophoresis, blotting and probing using radio-labelled *HRG1* DNA.
(TIFF)

S5 Fig. *RPP4-RPP5* structural diversity between Arabidopsis accessions analysed by Southern blotting and hybridization. Genomic DNA was isolated from the indicated accessions and digested with *Ase*I. DNA was separated using gel electrophoresis, blotting and probing using radio-labelled *RPP4* DNA.
(TIFF)

S6 Fig. *RPP1* structural diversity between Arabidopsis accessions analysed by Southern blotting and hybridization. Genomic DNA was isolated from the indicated accessions and digested with *Hae*II and *Pst*I. DNA was separated using gel electrophoresis, blotting and probing using radio-labelled *RPP1* DNA.
(TIFF)

S1 Table. *Arabidopsis thaliana* NBS-LRR gene family annotation. For each gene, chromosome, transcriptional start site (TSS) and termination site (TTS) are listed according to TAIR10 representative gene models. Historical recombination (cM/Mb) and diversity parameters (θ and π) were calculated for intragenic regions between TSS and TTS, using Eurasian and Swedish SNP datasets [72–74,82]. For *R* genes with known biological functions their given name is listed. In addition 3 columns have been added with TRUE/FALSE statements depending if the gene product has a match to the listed Pfam domain (TIR, NBS, LRR). Information is given on *R* gene locus structure, classified as either singleton, tandem, inverted or complex. The number of members within a repeated locus is listed.
(XLSX)

S2 Table. Crossover frequency within the *MRC1* NBS-LRR supercluster region. The table lists marker coordinates used to genotype double-selected *MRC1* crossover individuals, together with Col and Ler genotypes and interval lengths (bp). The number of crossovers identified in each interval is shown, together with cM/Mb. Eurasian and Swedish historical recombination rates estimated by LDhat are shown for the same intervals and NBS-LRR genes present in each interval are listed. A chi-square test was performed between the observed *MRC* crossover counts per interval and those expected at random using a 2×2 contingency table. $P<0.05$ values are listed in the *P* column, or listed as not significantly different (ns). The *P* adj. column shows the significance level after correction for multiple testing [78].
(DOCX)

S3 Table. Crossover frequency within the *MRC5* NBS-LRR supercluster region. The table lists marker coordinates used to genotype double-selected *MRC5* crossover individuals, together with Col and Ler genotypes and interval length (bp). The number of crossovers identified in each interval is shown, together with cM/Mb. Eurasian and Swedish historical recombination rates estimated by LDhat are shown for the same intervals and *R* genes present in each interval are listed. A chi-square test was performed between the observed crossover counts per interval and those expected at random using a 2×2 contingency table. $P<0.05$ values are listed in the *P* column, or listed as not significantly different (ns). The *P* adj. column shows the significance level after correction for multiple testing [78].
(DOCX)

S4 Table. Fine-mapping crossovers within the *HRG1 MRC1* map interval using dCAPs genotyping. The 'Genotyping Assay' column indicates whether a given marker coordinate was genotyped by KBiosciences (SNP), or via dCAPs assays.
(DOCX)

S5 Table. Fine-mapping crossovers within the *HRG2 HRG3 MRC5* map interval using dCAPs genotyping. The 'Genotyping Assay' column indicates whether a given marker coordinate was genotyped by KBiosciences (SNP), or via dCAPs assays.
(DOCX)

S6 Table. Fine-mapping crossovers within the *WRR4 MRC1* map interval using dCAPs genotyping. The 'Genotyping Assay' column indicates whether a given marker coordinate was genotyped by KBiosciences (SNP), or via dCAPs assays.
(DOCX)

S7 Table. Fine-mapping crossovers within the *CW9 MRC1* map interval using dCAPs genotyping. The 'Genotyping Assay' column indicates whether a given marker coordinate was genotyped by KBiosciences (SNP), or via dCAPs assays.
(DOCX)

S8 Table. Fine-mapping crossovers within the *HRG4-HRG5 MRC1* map interval using dCAPs genotyping. The 'Genotyping Assay' column indicates whether a given marker coordinate was genotyped by KBiosciences (SNP), or via dCAPs assays.
(DOCX)

S9 Table. Fine-mapping crossovers within the *HRG6 MRC1* map interval using dCAPs genotyping. The 'Genotyping Assay' column indicates whether a given marker coordinate was genotyped by KBiosciences (SNP), or via dCAPs assays.
(DOCX)

S10 Table. Fine-mapping crossovers within the *HRG7-HRG8 MRC5* map interval using dCAPs genotyping. The 'Genotyping Assay' column indicates whether a given marker coordinate was genotyped by KBiosciences (SNP), or via dCAPs assays.
(DOCX)

S11 Table. Fine-mapping crossovers within the *HRG9 MRC5* map interval using dCAPs genotyping. The 'Genotyping Assay' column indicates whether a given marker coordinate was genotyped by KBiosciences (SNP), or via dCAPs assays.
(DOCX)

S12 Table. Recombination rates of plant crossover hotspots. The mean and maximum crossover frequency (cM/Mb) measured at known plant hotspots is listed, together with the species studied, hotspot name, width of measured region, genetic distance (cM) and crossover frequency (cM/Mb), together with chromosome average recombination rates.
(DOCX)

S13 Table. Crossovers identified by genotyping by sequencing in a Col×Ler F_2 population. 192 Col×Ler F_2 individuals were analysed by genotyping-by-sequencing. The number of crossovers (CO) per F_2 and in total are listed, for the whole genome and for each chromosome separately.
(DOCX)

S14 Table. Measurement of *RAC1* recombination rate via pollen-typing. A 9.419 kb amplicon (Chr1: 11,288,146–11,297,565 bp) containing the *RAC1* gene was amplified using allele-specific oligonucleotides to estimate the number of parental (non-recombinant) and crossover molecules per μl of Col/Ler F_1 pollen genomic DNA. Genetic distance was calculated as described [83].
(DOCX)

S15 Table. Crossover distributions across the *RAC1 R* gene hotspot analysed via pollen typing. 181 single *RAC1* crossover molecules were Sanger sequenced to identify recombination sites to the resolution of single polymorphisms. Col and Ler genotypes and number of crossovers per interval are listed, together with cM/Mb.
(DOCX)

S16 Table. *RAC1* crossover read pair analysis pipeline. The number of read pairs surviving sequential analysis filters are listed in order to identify *RAC1* crossover read pairs. Paired end reads (1 and 2) were separated and aligned to the Col or Ler *RAC1* parental template sequences, allowing only exact matches (Mapped). Read pair ends that mapped to both Col and Ler were then excluded (Unique). Read pair ends (1 and 2) that mapped to Col and Ler were identified (Matched), where the Ler mapping read had a lower coordinate than the Col mapping read (Orientate), and that were on opposite strands (Strand). These filters yielded 182,909 crossover read pairs.
(DOCX)

S17 Table. CTT-repeat motifs associated with high and low *MRC* recombination NBS-LRR genes. NBS-LRR genes located within the *MRC* genetic maps were divided into two groups according to whether their interval had higher or lower crossover frequency (cM/Mb) compared with the male Col×Ler genome average (4.82 cM/Mb) [71]. We then matched the position weight matrix of a CTT-repeat motif previously identified as enriched at Arabidopsis historical crossover hotspots (CTTCTTCTTCTTCTTC) [37] to +1 kb windows around NBS-LRR gene transcriptional start sites (TSSs), allowing matches with >80% identity. The location, width and sequence of matching CTT motifs are listed. Motif coordinates are given relative to the 1 kb of sequence matched to.
(DOCX)

S18 Table. Measurement of *I5a* genetic distance using fluorescent pollen and flow cytometry. The first column lists the accession the *I5a* reporter (Col-0) was crossed to. In the case of Col-0 this represents data from Col-0/Col-0 homozygotes. The total number of pollen counted is listed, in addition to the number with red fluorescence alone (Red), yellow alone (yellow), both colors (Red+Yellow), or neither fluorescence (None). Genetic distance is calculated as cM = 100 × (Y/(Y+RY)). Where Y is the number of yellow alone pollen grains and RG is the number of both yellow and red fluorescent pollen grains. To test whether recombinant and non-recombinant counts were significantly different between replicate groups we used a generalized linear model (GLM), assuming that the count data is binomially distributed. Tests were performed between F_1 genotypes and Col/Col homozygotes.
(DOCX)

S19 Table. *RAC1* primer sequences used for allele-specific PCR amplification. Positions annealing to Col/Ler polymorphic bases in the allele-specific primers are indicated in red. The accession each primer anneals to is listed (Col or Ler), together with a code indicating whether the primer is forward or reverse (F or R) and used in the first or second round of allele-specific amplification (1 or 2).
(DOCX)

Acknowledgments

We kindly thank Detlef Weigel and Magnus Nordborg for providing 1,001 Genomes Project SNP data used for LDhat analysis.

Author Contributions

Conceived and designed the experiments: KC CR HS PAZ CJU NEY CG IRH. Performed the experiments: KC CR HS PAZ CJU CG. Analyzed the data: KC CR HS PAZ CJU XZ TJH NEY CG MJ CM GM GPC IRH. Contributed reagents/materials/analysis tools: CM GM GPC. Wrote the paper: KC CR HS PAZ CJU XZ TJH NEY CG MJ CM GM GPC IRH.

References

1. Villeneuve AM, Hillers KJ. Whence meiosis? Cell. 2001; 106: 647–50. PMID: 11572770
2. Mercier R, Mézard C, Jenczewski E, Macaisne N, Grelon M. The Molecular Biology of Meiosis in Plants. Annu Rev Plant Biol. 2015; 66:297–327. doi: 10.1146/annurev-arplant-050213-035923 PMID: 25494464
3. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. Cell. 1983; 33: 25–35. PMID: 6380756
4. Lichten M, Goldman AS. Meiotic recombination hotspots. Annu Rev Genet. 1995; 29: 423–44. PMID: 8825482

5. De Massy B. Initiation of meiotic recombination: how and where? Conservation and specificities among eukaryotes. Annu Rev Genet. 2013; 47: 563–99. doi: 10.1146/annurev-genet-110711-155423 PMID: 24050176
6. Kauppi L, Jeffreys AJ, Keeney S. Where the crossovers are: recombination distributions in mammals. Nat Rev Genet. 2004; 5: 413–24. PMID: 15153994
7. Choi K, Henderson IR. Meiotic recombination hotspots—a comparative view. Plant J. 2015; 83: 52–61. doi: 10.1111/tpj.12870 PMID: 25925869
8. Haldane J. Disease and evolution. Ric Sci Suppl A. 1949; 19: 68–76.
9. Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, Cano LM, et al. Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. Nature. 2009; 461: 393–8. doi: 10.1038/nature08358 PMID: 19741609
10. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449: 851–61. PMID: 17943122
11. Kämper J, Kahmann R, Bölker M, Ma L-J, Brefort T, Saville BJ, et al. Insights from the genome of the biotrophic fungal plant pathogen Ustilago maydis. Nature. 2006; 444: 97–101. PMID: 17080091
12. Dong S, Raffaele S, Kamoun S. The two-speed genomes of filamentous pathogens: waltz with plants. Curr Opin Genet Dev. 2015; 35: 57–65. doi: 10.1016/j.gde.2015.09.001 PMID: 26451981
13. Jones JDG, Dangl JL. The plant immune system. Nature. 2006; 444: 323–329. PMID: 17108957
14. Chisholm ST, Coaker G, Day B, Staskawicz BJ. Host-microbe interactions: shaping the evolution of the plant immune response. Cell. 2006; 124: 803–14. PMID: 16497589
15. Zipfel C. Early molecular events in PAMP-triggered immunity. Curr Opin Plant Biol. 2009; 12: 414–20. doi: 10.1016/j.pbi.2009.06.003 PMID: 19608450
16. Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell. 2003; 15: 809–34. PMID: 12671079
17. Hofberger JA, Zhou B, Tang H, Jones JDG, Schranz ME. A novel approach for multi-domain and multi-gene family identification provides insights into evolutionary dynamics of disease resistance genes in core eudicot plants. BMC Genomics. 2014; 15: 966. doi: 10.1186/1471-2164-15-966 PMID: 25380807
18. Deslandes L, Olivier J, Peeters N, Feng DX, Khounlotham M, Boucher C, et al. Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. Proc Natl Acad Sci U S A. 2003; 100: 8024–9. PMID: 12788974
19. Kim MG, da Cunha L, McFall AJ, Belkhadir Y, DebRoy S, Dangl JL, et al. Two Pseudomonas syringae type III effectors inhibit RIN4-regulated basal defense in Arabidopsis. Cell. 2005; 121: 749–59. PMID: 15935761
20. Van der Hoorn RAL, Kamoun S. From Guard to Decoy: a new model for perception of plant pathogen effectors. Plant Cell. 2008; 20: 2009–17. doi: 10.1105/tpc.108.060194 PMID: 18723576
21. Zhou J-M, Chai J. Plant pathogenic bacterial type III effectors subdue host responses. Curr Opin Microbiol. 2008; 11: 179–85. doi: 10.1016/j.mib.2008.02.004 PMID: 18372208
22. Zipfel C, Rathjen JP. Plant immunity: AvrPto targets the frontline. Curr Biol. 2008; 18: R218–20. doi: 10.1016/j.cub.2008.01.016 PMID: 18334200
23. Dodds PN, Lawrence GJ, Catanzariti A-M, Teh T, Wang C-IA, Ayliffe MA, et al. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. Proc Natl Acad Sci U S A. 2006; 103: 8888–93. PMID: 16731621
24. Gao L, Roux F, Bergelson J. Quantitative fitness effects of infection in a gene-for-gene system. New Phytol. 2009; 184: 485–94. doi: 10.1111/j.1469-8137.2009.02959.x PMID: 19659661
25. Roux F, Gao L, Bergelson J. Impact of initial pathogen density on resistance and tolerance in a polymorphic disease resistance gene system in Arabidopsis thaliana. Genetics. 2010; 185: 283–91. doi: 10.1534/genetics.109.112383 PMID: 20142437
26. Karasov TL, Kniskern JM, Gao L, DeYoung BJ, Ding J, Dubiella U, et al. The long-term maintenance of a resistance polymorphism through diffuse interactions. Nature. 2014; 512: 436–440. doi: 10.1038/nature13439 PMID: 25043057
27. Tian D, Traw MB, Chen JQ, Kreitman M, Bergelson J. Fitness costs of R-gene-mediated resistance in Arabidopsis thaliana. Nature. 2003; 423: 74–77. PMID: 12721627
28. Narusaka M, Shirasu K, Noutoshi Y, Kubo Y, Shiraishi T, Iwabuchi M, et al. RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. Plant J. 2009; 60: 218–26. doi: 10.1111/j.1365-313X.2009.03949.x PMID: 19519800

29. Le Roux C, Huet G, Jauneau A, Camborde L, Trémousaygue D, Kraut A, et al. A Receptor Pair with an Integrated Decoy Converts Pathogen Disabling of Transcription Factors to Immunity. Cell. 2015; 161: 1074–1088. doi: 10.1016/j.cell.2015.04.025 PMID: 26000483

30. Sarris PF, Duxbury Z, Huh SU, Ma Y, Segonzac C, Sklenar J, et al. A Plant Immune Receptor Detects Pathogen Effectors that Target WRKY Transcription Factors. Cell. 2015; 161: 1089–1100. doi: 10.1016/j.cell.2015.04.024 PMID: 26000484

31. Saucet SB, Ma Y, Sarris PF, Furzer OJ, Sohn KH, Jones JDG. Two linked pairs of Arabidopsis TNL resistance genes independently confer recognition of bacterial effector AvrRps4. Nat Commun. 2015; 6: 6338. doi: 10.1038/ncomms7338 PMID: 25744164

32. Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL, et al. Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. PLoS Biol. 2007; 5: e236. PMID: 17803357

33. Chae E, Bomblies K, Kim S-T, Karelina D, Zaidem M, Ossowski S, et al. Species-wide Genetic Incompatibility Analysis Identifies Immune Genes as Hot Spots of Deleterious Epistasis. Cell. 2014; 159: 1341–51. doi: 10.1016/j.cell.2014.10.049 PMID: 25467443

34. Jeuken MJW, Zhang NW, McHale LK, Pelgrom K, den Boer E, Lindhout P, et al. Rin4 causes hybrid necrosis and race-specific resistance in an interspecific lettuce hybrid. Plant Cell. 2009; 21: 3368–78. doi: 10.1105/tpc.109.070334 PMID: 19855048

35. Guo Y-L, Fitz J, Schneeberger K, Ossowski S, Cao J, Weigel D. Genome-wide comparison of nucleotide-binding site-leucine-rich repeat-encoding genes in Arabidopsis. Plant Physiol. 2011; 157: 757–69. doi: 10.1104/pp.111.181990 PMID: 21810963

36. Hu TT, Pattyn P, Bakker EG, Cao J, Cheng J-F, Clark RM, et al. The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. Nat Genet; 2011; 43: 476–81. doi: 10.1038/ng.807 PMID: 21478890

37. Choi K, Zhao X, Kelly KA, Venn O, Higgins JD, Yelina NE, et al. Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. Nat Genet. 2013; 45: 1327–36. doi: 10.1038/ng.2766 PMID: 24056716

38. Paape T, Zhou P, Branca A, Briskine R, Young N, Tiffin P. Fine-scale population recombination rates, hotspots, and correlates of recombination in the Medicago truncatula genome. Genome Biol Evol. 2012; 4: 726–37. doi: 10.1093/gbe/evs046 PMID: 22554552

39. Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, et al. Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel. Nat Genet. 2012; 44: 212–6. doi: 10.1038/ng.1042 PMID: 22231484

40. Rose LE, Bittner-Eddy PD, Langley CH, Holub EB, Michelmore RW, Beynon JL. The maintenance of extreme amino acid diversity at the disease resistance gene, RPP13, in Arabidopsis thaliana. Genetics. 2004; 166: 1517–27. PMID: 15082565

41. Gos G, Slotte T, Wright SI. Signatures of balancing selection are maintained at disease resistance loci following mating system evolution and a population bottleneck in the genus Capsella. BMC Evol Biol. 2012; 12: 152. doi: 10.1186/1471-2148-12-152 PMID: 22909344

42. Tian D, Araki H, Stahl E, Bergelson J, Kreitman M. Signature of balancing selection in Arabidopsis. Proc Natl Acad Sci U S A. 2002; 99: 11525–30. PMID: 12172007

43. Allen RL, Bittner-Eddy PD, Grenville-Briggs LJ, Meitz JC, Rehmany AP, Rose LE, et al. Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. Science. 2004; 306: 1957–60. PMID: 15591208

44. Bakker EG, Toomajian C, Kreitman M, Bergelson J. A genome-wide survey of R gene polymorphisms in Arabidopsis. Plant Cell. 2006; 18: 1803–18. PMID: 16798885

45. Bergelson J, Kreitman M, Stahl EA, Tian D. Evolutionary dynamics of plant R-genes. Science (80-). 2001; 292: 2281–2285.

46. Mondragón-Palomino M, Meyers BC, Michelmore RW, Gaut BS. Patterns of positive selection in the complete NBS-LRR gene family of Arabidopsis thaliana. Genome Res. 2002; 12: 1305–15. PMID: 12213767

47. Tellier A, Brown JKM. Stability of genetic polymorphism in host-parasite interactions. Proc Biol Sci. 2007; 274: 809–17. PMID: 17251091

48. Delaney DE, Webb CA, Hulbert SH. A Novel Rust Resistance Gene in Maize Showing Overdominance. The American Phytopathological Society; 2007 MPMI.1998.11.3.242

49. Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, et al. Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the Cf-4/9 locus of tomato. Cell. 1997; 91: 821–832. PMID: 9413991

50. McDowell JM, Dhandaydham M, Long TA, Aarts MG, Goff S, Holub EB, et al. Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. Plant Cell. 1998; 10: 1861–74. PMID: 9811794

51. Chin DB, Arroyo-Garcia R, Ochoa OE, Kesseli R V, Lavelle DO, Michelmore RW. Recombination and spontaneous mutation at the major cluster of resistance genes in lettuce (Lactuca sativa). Genetics. 2001; 157: 831–49. PMID: 11157000

52. Noël L, Moores TL, van Der Biezen EA, Parniske M, Daniels MJ, Parker JE, et al. Pronounced intraspecific haplotype divergence at the RPP5 complex disease resistance locus of Arabidopsis. Plant Cell. 1999; 11: 2099–112. PMID: 10559437

53. Parker JE, Coleman MJ, Szabò V, Frost LN, Schmidt R, van der Biezen EA, et al. The Arabidopsis downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6. Plant Cell. 1997; 9: 879–94. PMID: 9212464

54. Richter TE, Pryor TJ, Bennetzen JL, Hulbert SH. New rust resistance specificities associated with recombination in the Rp1 complex in maize. Genetics. 1995; 141: 373–381. PMID: 8536984

55. Sun Q, Collins NC, Ayliffe M, Smith SM, Drake J, Pryor T, et al. Recombination between paralogues at the Rp1 rust resistance locus in maize. Genetics. 2001; 158: 423–438. PMID: 11333250

56. Sudupak MA, Bennetzen JL, Hulbert SH. Unequal exchange and meiotic instability of disease-resistance genes in the Rp1 region of maize. Genetics. 1993; 133: 119–25. PMID: 8417982

57. Ellis JG, Lawrence GJ, Finnegan EJ, Anderson PA. Contrasting complexity of two rust resistance loci in flax. Proc Natl Acad Sci U S A. 1995; 92: 4185–8. PMID: 11607543

58. Luck JE, Lawrence GJ, Dodds PN, Shepherd KW, Ellis JG. Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. Plant Cell. 2000; 12: 1367–77. PMID: 10948256

59. Lawrence GJ, Anderson PA, Dodds PN, Ellis JG. Relationships between rust resistance genes at the M locus in flax. Mol Plant Pathol. 2010; 11: 19–32. doi: 10.1111/j.1364-3703.2009.00563.x PMID: 20078773

60. Borhan MH, Holub EB, Beynon JL, Rozwadowski K, Rimmer SR. The arabidopsis TIR-NB-LRR gene RAC1 confers resistance to Albugo candida (white rust) and is dependent on EDS1 but not PAD4. Mol Plant Microbe Interact. 2004; 17: 711–9. PMID: 15242165

61. Speulman E, Bouchez D, Holub EB, Beynon JL. Disease resistance gene homologs correlate with disease resistance loci of Arabidopsis thaliana. Plant J. 1998; 14: 467–74. PMID: 9670562

62. Borhan MH, Brose E, Beynon JL, Holub EB. White rust (Albugo candida) resistance loci on three Arabidopsis chromosomes are closely linked to downy mildew (Peronospora parasitica) resistance loci. Mol Plant Pathol. 2001; 2: 87–95. PMID: 20572995

63. Shilo S, Melamed-Bessudo C, Dorone Y, Barkai N, Levy AA. DNA Crossover Motifs Associated with Epigenetic Modifications Delineate Open Chromatin Regions in Arabidopsis. Plant Cell. 2015; 27.

64. Wijnker E, Velikkakam James G, Ding J, Becker F, Klasen JR, Rawat V, et al. The genomic landscape of meiotic crossovers and gene conversions in Arabidopsis thaliana. Elife. 2013; 2: e01426. doi: 10.7554/eLife.01426 PMID: 24347547

65. Francis KE, Lam SY, Harrison BD, Bey AL, Berchowitz LE, Copenhaver GP. Pollen tetrad-based visual assay for meiotic recombination in Arabidopsis. Proc Natl Acad Sci U S A. 2007; 104: 3913–3918. PMID: 17360452

66. Berchowitz LE, Copenhaver GP. Fluorescent Arabidopsis tetrads: a visual assay for quickly developing large crossover and crossover interference data sets. Nat Protoc. 2008; 3: 41–50. doi: 10.1038/nprot.2007.491 PMID: 18193020

67. Yelina NE, Ziolkowski PA, Miller N, Zhao X, Kelly KA, Muñoz DF, et al. High-throughput analysis of meiotic crossover frequency and interference via flow cytometry of fluorescent pollen in Arabidopsis thaliana. Nat Protoc. 2013; 8: 2119–2134. doi: 10.1038/nprot.2013.131 PMID: 24113785

68. Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, et al. A high-throughput Arabidopsis reverse genetics system. Plant Cell. 2002; 14: 2985–94. PMID: 12468722

69. Parinov S, Sevugan M, Ye D, Yang WC, Kumaran M, Sundaresan V. Analysis of flanking sequences from dissociation insertion lines: a database for reverse genetics in Arabidopsis. Plant Cell. 1999; 11: 2263–70. PMID: 10590156

70. Neff MM, Neff JD, Chory J, Pepper AE. dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in Arabidopsis thaliana genetics. Plant J. 1998; 14: 387–92. PMID: 9628033

71. Giraut L, Falque M, Drouaud J, Pereira L, Martin OC, Mézard C. Genome-wide crossover distribution in Arabidopsis thaliana meiosis reveals sex-specific patterns along chromosomes. PLoS Genet. 2011; 7: e1002354. doi: 10.1371/journal.pgen.1002354 PMID: 22072983

72. Long Q, Rabanal FA, Meng D, Huber CD, Farlow A, Platzer A, et al. Massive genomic variation and strong selection in Arabidopsis thaliana lines from Sweden. Nat Genet. 2013; 45: 884–90. doi: 10.1038/ng.2678 PMID: 23793030

73. Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, et al. Whole-genome sequencing of multiple Arabidopsis thaliana populations. Nat Genet. 2011; 43: 956–63. doi: 10.1038/ng.911 PMID: 21874002

74. Auton A, McVean G. Recombination rate estimation in the presence of hotspots. Genome Res. 2007; 17: 1219–27. PMID: 17623807

75. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature. 2000; 408: 796–815. PMID: 11130711

76. Botella MA, Parker JE, Frost LN, Bittner-Eddy PD, Beynon JL, Daniels MJ, et al. Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct Peronospora parasitica avirulence determinants. Plant Cell. 1998; 10: 1847–60. PMID: 9811793

77. Kim SH, Kwon S Il, Saha D, Anyanwu NC, Gassmann W. Resistance to the Pseudomonas syringae effector HopA1 is governed by the TIR-NBS-LRR protein RPS6 and is enhanced by mutations in SRFR1. Plant Physiol. 2009; 150: 1723–32. doi: 10.1104/pp.109.139238 PMID: 19525323

78. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 1995; 57: 289–300.

79. Rowan BA, Patel V, Weigel D, Schneeberger K. Rapid and Inexpensive Whole-Genome Genotyping-by-Sequencing for Crossover Localization and Fine-Scale Genetic Mapping. G3 (Bethesda). 2015; 5: 385–98.

80. Yelina NE, Lambing C, Hardcastle TJ, Zhao X, Santos B, Henderson IR. DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in Arabidopsis. Genes Dev. 2015; 29: 2183–202. doi: 10.1101/gad.270876.115 PMID: 26494791

81. Auton A, McVean G. Estimating recombination rates from genetic variation in humans. Methods Mol Biol. 2012; 856: 217–37. doi: 10.1007/978-1-61779-585-5_9 PMID: 22399461

82. Charlesworth B, Charlesworth D. Elements of Evolutionary Genetics. Roberts and Company Publishers.; 2010.

83. Drouaud J, Mézard C. Characterization of meiotic crossovers in pollen from Arabidopsis thaliana. Methods Mol Biol. 2011; 745: 223–49. doi: 10.1007/978-1-61779-129-1_14 PMID: 21660698

84. Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, et al. Epigenetic remodeling of meiotic crossover frequency in Arabidopsis thaliana DNA methyltransferase mutants. PLoS Genet. 2012; 8: e1002844. doi: 10.1371/journal.pgen.1002844 PMID: 22876192

85. Kauppi L, May CA, Jeffreys AJ. Analysis of meiotic recombination products from human sperm. Keeney S, editor. Methods in molecular biology (Clifton, N.J.). Totowa, NJ: Humana Press; 2009. doi: 10.1007/978-1-59745-527-5

86. Arbeithuber B, Betancourt AJ, Ebner T, Tiemann-Boege I. Crossovers are associated with mutation and biased gene conversion at recombination hotspots. Proc Natl Acad Sci U S A. 2015; 112: 2109–14. doi: 10.1073/pnas.1416622112 PMID: 25646453

87. Hellsten U, Wright KM, Jenkins J, Shu S, Yuan Y, Wessler SR, et al. Fine-scale variation in meiotic recombination in Mimulus inferred from population shotgun sequencing. Proc Natl Acad Sci U S A. 2013; 110: 19478–82. doi: 10.1073/pnas.1319032110 PMID: 24225854

88. Williams SJ, Sohn KH, Wan L, Bernoux M, Sarris PF, Segonzac C, et al. Structural basis for assembly and function of a heterodimeric plant immune receptor. Science. 2014; 344: 299–303. doi: 10.1126/science.1247357 PMID: 24744375

89. Harris CJ, Slootweg EJ, Goverse A, Baulcombe DC. Stepwise artificial evolution of a plant disease resistance gene. Proc Natl Acad Sci U S A. 2013; 110: 21189–94. doi: 10.1073/pnas.1311134110 PMID: 24324167

90. Steinbrenner AD, Goritschnig S, Staskawicz BJ. Recognition and activation domains contribute to allele-specific responses of an Arabidopsis NLR receptor to an oomycete effector protein. PLoS Pathog. 2015; 11: e1004665. doi: 10.1371/journal.ppat.1004665 PMID: 25671309

91. Ravensdale M, Bernoux M, Ve T, Kobe B, Thrall PH, Ellis JG, et al. Intramolecular interaction influences binding of the Flax L5 and L6 resistance proteins to their AvrL567 ligands. PLoS Pathog. 2012; 8: e1003004. doi: 10.1371/journal.ppat.1003004 PMID: 23209402

92. Chodavarapu RK, Feng S, Bernatavichute Y V, Chen P-Y, Stroud H, Yu Y, et al. Relationship between nucleosome positioning and DNA methylation. Nature. 2010; 466: 388–392. doi: 10.1038/nature09147 PMID: 20512117

93. Flores O, Orozco M. nucleR: a package for non-parametric nucleosome positioning. Bioinformatics. 2011; 27: 2149–50. doi: 10.1093/bioinformatics/btr345 PMID: 21653521

94. Berlin K, Koren S, Chin C-S, Drake JP, Landolin JM, Phillippy AM. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. Nat Biotechnol. 2015; doi: 10.1038/nbt.3238

95. Borts RH, Haber JE. Meiotic recombination in yeast: alteration by multiple heterozygosities. Science. 1987; 237: 1459–65. PMID: 2820060

96. Dooner HK. Genetic Fine Structure of the BRONZE Locus in Maize. Genetics. 1986; 113: 1021–36. PMID: 17246338

97. Ziolkowski PA, Berchowitz LE, Lambing C, Yelina NE, Zhao X, Kelly KA, et al. Juxtaposition of heterozygosity and homozygosity during meiosis causes reciprocal crossover remodeling via interference. Elife. 2015; 4: e03708.

98. Timmermans MC, Das OP, Bradeen JM, Messing J. Region-specific cis- and trans-acting factors contribute to genetic variability in meiotic recombination in maize. Genetics. 1997; 146: 1101–1113. PMID: 9215911

99. Esch E, Szymaniak JM, Yates H, Pawlowski WP, Buckler ES. Using crossover breakpoints in recombinant inbred lines to identify quantitative trait loci controlling the global recombination frequency. Genetics. 2007; 177: 1851–1858. PMID: 17947409

100. Yao H, Schnable PS. Cis-effects on meiotic recombination across distinct a1-sh2 intervals in a common Zea genetic background. Genetics. 2005; 170: 1929–1944. PMID: 15937141

101. Barth S, Melchinger AE, Devezi-Savula B, Lübberstedt T. Influence of genetic background and heterozygosity on meiotic recombination in Arabidopsis thaliana. Genome. 2001; 44: 971–978. PMID: 11768224

102. López E, Pradillo M, Oliver C, Romero C, Cuñado N, Santos JL. Looking for natural variation in chiasma frequency in Arabidopsis thaliana. J Exp Bot. 2012; 63: 887–94. doi: 10.1093/jxb/err319 PMID: 22048037

103. Salomé PA, Bomblies K, Fitz J, Laitinen RAE, Warthmann N, Yant L, et al. The recombination landscape in Arabidopsis thaliana F2 populations. Heredity (Edinb). 2012; 108: 447–55.

104. Bauer E, Falque M, Walter H, Bauland C, Camisan C, Campo L, et al. Intraspecific variation of recombination rate in maize. Genome Biol. 2013; 14: R103. PMID: 24050704

105. McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, Sun Q, et al. Genetic properties of the maize nested association mapping population. Science. 2009; 325: 737–740. doi: 10.1126/science.1174320 PMID: 19661427

106. Kuang H, Caldwell KS, Meyers BC, Michelmore RW. Frequent sequence exchanges between homologs of RPP8 in Arabidopsis are not necessarily associated with genomic proximity. Plant J. 2008; 54: 69–80. doi: 10.1111/j.1365-313X.2008.03408.x PMID: 18182023

107. Kuang H, Woo S-S, Meyers BC, Nevo E, Michelmore RW. Multiple genetic processes result in heterogeneous rates of evolution within the major cluster disease resistance genes in lettuce. Plant Cell. 2004; 16: 2870–94. PMID: 15494555

108. Michelmore RW, Meyers BC. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res. 1998; 8: 1113–30. PMID: 9847076

109. Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, et al. The scale of population structure in Arabidopsis thaliana. PLoS Genet. 2010; 6: e1000843. doi: 10.1371/journal.pgen.1000843 PMID: 20169178

110. Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, et al. Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of Arabidopsis thaliana. PLoS Genet. 2010; 6: e1000890. doi: 10.1371/journal.pgen.1000890 PMID: 20361058

111. Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, et al. Recombination and linkage disequilibrium in Arabidopsis thaliana. Nat Genet. 2007; 39: 1151–1155. PMID: 17676040

112. Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, et al. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science. 2010; 327: 836–40. doi: 10.1126/science.1183439 PMID: 20044539

113. Grey C, Barthès P, Chauveau-Le Friec G, Langa F, Baudat F, de Massy B. Mouse PRDM9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination. PLoS Biol. 2011; 9: e1001176. doi: 10.1371/journal.pbio.1001176 PMID: 22028627

114. Baker CL, Walker M, Kajita S, Petkov PM, Paigen K. PRDM9 binding organizes hotspot nucleosomes and limits Holliday junction migration. Genome Res. 2014; 24: 724–732. doi: 10.1101/gr.170167.113 PMID: 24604780

115. Brick K, Smagulova F, Khil P, Camerini-Otero RD, Petukhova G V. Genetic recombination is directed away from functional genomic elements in mice. Nature. 2012; 485: 642–5. doi: 10.1038/nature11089 PMID: 22660327

116. Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C, MacFie TS, et al. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science. 2010; 327: 876–9. doi: 10.1126/science.1182363 PMID: 20044541

117. Parvanov ED, Petkov PM, Paigen K. Prdm9 controls activation of mammalian recombination hotspots. Science. 2010; 327: 835. doi: 10.1126/science.1181495 PMID: 20044538

118. Copenhaver GP, Nickel K, Kuromori T, Benito MI, Kaul S, Lin X, et al. Genetic definition and sequence analysis of Arabidopsis centromeres. Science. 1999; 286: 2468–74. PMID: 10617454

119. McMahill MS, Sham CW, Bishop DK. Synthesis-dependent strand annealing in meiosis. PLoS Biol. 2007; 5: e299. PMID: 17988174

120. Martini E, Borde V, Legendre M, Audic S, Regnault B, Soubigou G, et al. Genome-wide analysis of heteroduplex DNA in mismatch repair-deficient yeast cells reveals novel properties of meiotic recombination pathways. PLoS Genet. 2011; 7: e1002305. doi: 10.1371/journal.pgen.1002305 PMID: 21980306

121. Parniske M, Jones JDG. Recombination between diverged clusters of the tomato Cf-9 plant disease resistance gene family. Proc Natl Acad Sci U S A. 1999; 96: 5850–5. PMID: 10318973

122. Song WY, Pi LY, Wang GL, Gardner J, Holsten T, Ronald PC. Evolution of the rice Xa21 disease resistance gene family. Plant Cell. 1997; 9: 1279–87. PMID: 9286106

123. Meyers BC, Chin DB, Shen KA, Sivaramakrishnan S, Lavelle DO, Zhang Z, et al. The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. Plant Cell. 1998; 10: 1817–32. PMID: 9811791

124. Dodds PN, Lawrence GJ, Ellis JG. Contrasting modes of evolution acting on the complex N locus for rust resistance in flax. Plant J. 2001; 27: 439–53. PMID: 11576428

125. Anderson PA, Lawrence GJ, Morrish BC, Ayliffe MA, Finnegan EJ, Ellis JG. Inactivation of the flax rust resistance gene M associated with loss of a repeated unit within the leucine-rich repeat coding region. Plant Cell. 1997; 9: 641–51. PMID: 9144966

126. Wulff BBH, Thomas CM, Parniske M, Jones JDG. Genetic variation at the tomato Cf-4/Cf-9 locus induced by EMS mutagenesis and intralocus recombination. Genetics. 2004; 167: 459–470. PMID: 15166169

127. Hulbert SH, Bennetzen JL. Recombination at the Rp1 locus of maize. Mol Gen Genet. 1991; 226: 377–382. PMID: 1674815

128. Moffett P, Farnham G, Peart J, Baulcombe DC. Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. EMBO J. 2002; 21: 4511–9. PMID: 12198153

129. Trowsdale J, Knight JC. Major histocompatibility complex genomics and human disease. Annu Rev Genomics Hum Genet. 2013; 14: 301–23. doi: 10.1146/annurev-genom-091212-153455 PMID: 23875801

130. Stahl EA, Dwyer G, Mauricio R, Kreitman M, Bergelson J. Dynamics of disease resistance polymorphism at the Rpm1 locus of Arabidopsis. Nature. 1999; 400: 667–71. PMID: 10458161

131. Jeffreys AJ, Kauppi L, Neumann R. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nat Genet. 2001; 29: 217–22. PMID: 11586303

132. Raffaele S, Kamoun S. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Microbiol.; 2012; 10: 417–30. doi: 10.1038/nrmicro2790 PMID: 22565130

133. Kyes SA, Kraemer SM, Smith JD. Antigenic variation in Plasmodium falciparum: gene organization and regulation of the var multigene family. Eukaryot Cell. 2007; 6: 1511–20. PMID: 17644655

134. Glover L, Hutchinson S, Alsford S, McCulloch R, Field MC, Horn D. Antigenic variation in African trypanosomes: the importance of chromosomal and nuclear context in VSG expression control. Cell Microbiol. 2013; 15: 1984–93. doi: 10.1111/cmi.12215 PMID: 24047558

135. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 2010; 59: 307–21. doi: 10.1093/sysbio/syq010 PMID: 20525638

136. Drouaud J, Khademian H, Giraut L, Zanni V, Bellalou S, Henderson IR, et al. Contrasted patterns of crossover and non-crossover at Arabidopsis thaliana meiotic recombination hotspots. PLoS Genet. 2013; 9: e1003922. doi: 10.1371/journal.pgen.1003922 PMID: 24244190

137. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9: 357–9. doi: 10.1038/nmeth.1923 PMID: 22388286