Recurrent chromosome destabilization through repeat-mediated rearrangements in a fungal pathogen

Simone Fouché[1,2,#], Ursula Oggenfuss[1,#], Bruce A. McDonald[2], Daniel Croll[1,*]

[1] Laboratory of Evolutionary Genetics, Institute of Biology, University of Neuchâtel, CH-2000, Neuchâtel, Switzerland

[2] Plant Pathology, Institute of Integrative Biology, ETH Zurich, CH-8092 Zurich, Switzerland

[*] Corresponding author. Email: daniel.croll@unine.ch

[#] Co-first authors

Abstract

Genomic instability caused by chromosomal rearrangements has severe consequences for organismal fitness and progression of cancerous cell lines. The triggers of destabilized chromosomes remain poorly understood but likely co-locate with fragile sites. Here, we retrace a runaway chromosomal degeneration process observed in a fungal pathogen using telomere-to-telomere assemblies across an experimental progeny. We show that the same fragile sites triggered reproducible, large-scale rearrangements through non-allelic recombination. Across our four-generation progeny, chromosomal rearrangements were accompanied by non-disjunction events leading to aneuploid progeny with up to four chromosomal copies. We identify a specific transposable element family co-locating with fragile sites, likely triggering ongoing repeated chromosomal degeneration. The element has recently been associated with lower virulence of the pathogen and has undergone an expansion of copy numbers across the genome. Chromosome sequences are also targeted by repeat-induced point mutation, a genome defense mechanism actively leading to hypermutation on duplicated sequences. Our work identifies the exact sequence triggers that initiate chromosome instability and perpetuate degenerative cycles. Dissecting proximate causes leading to runaway chromosomal degeneration could expand our understanding of chromosomal evolution beyond fungal pathogens.

Introduction

Meiosis is a highly conserved process in eukaryotes, whereby homologous chromosomes pair, undergo recombination, and separate into daughter cells. Aberrations during the faithful transmission of chromosomes through meiosis can have serious consequences for an organism. Non-disjunction events resulting in additional or fewer chromosomal copies occur frequently in humans, and are the leading causes of miscarriage (*1*). Similarly, non-disjunction events and chromosome rearrangements occur in somatic lines, and may lead to genomic instability, which is a hallmark of cancers (*2*). However, factors that cause fragility of pre-cancerous genomes or that trigger meiotic errors are largely unknown. Chromosomal breakage is a major factor contributing to instability and occurs often at specific loci referred to as fragile sites (*3*). Though the locations of many mitotic fragile sites have been identified in the human genome (*4*), how they contribute mechanistically to breakage remains poorly understood. Fragile sites for non-allelic homologous or ectopic recombination during meiosis remain largely unknown, possibly because major rearrangements tend to be lethal and are quickly purged from the population by purifying selection. Investigations of non-lethal rearrangements offer a promising approach to unravel sequence determinants of fragile sites.

Accessory chromosomes in plant and fungal genomes are powerful models to study non-lethal chromosome rearrangements. These chromosomes (also called supernumerary or B) are present in a karyotype in addition to the regular chromosomes, contain no essential genes, and show broad presence/absence within species (*5*). It is unclear what mechanisms generate these chromosome rearrangements and whether fragile sites are involved. For example, the accessory chromosomes of *Magnaporthe oryzae* contain many rearrangements and frequent interchromosome translocations between core and accessory chromosomes (*6*). Accessory chromosomes are hypothesized to originate from core chromosomes through rearrangements involving the terminal regions of core chromosomes. *Zymoseptoria tritici*, a haploid fungal pathogen of wheat, exhibits some of the most extreme degrees of structural variation observed within fungal species, including large differences in terms of chromosome length, transposable element (TE) and gene content, and recombination rate, as well as telomere and centromere structures (*7–13*). Chromosome rearrangements occur frequently during meiosis, with higher rates of rearrangements observed in the accessory chromosomes (*7*, *14*).

We analyzed a rearranged accessory chromosome of *Z. tritici*, , which was generated repeatedly during meiosis through an unknown mechanism (*7*). Here, we show how the rearranged chromosome was generated and how further rearrangements occurred through subsequent rounds of meiosis. We also find

67 evidence that a transposable element acts as the specific sequence trigger of the chromosome rearrangement
68 and degenerative cycles.
69
70
71 **Results**
72 ***Origin of a complex chromosome rearrangement***
73 The rearrangement of the *Z. tritici* accessory chromosome 17 was first discovered in two F1 progeny A66.2
74 and A2.2 (fig. 1A) (*7*). The enlarged nature of chromosome 17 in these progeny was confirmed by pulsed-
75 field gel electrophoresis (PFGE) as well as southern hybridization (fig. 1B) (*7*). To assess how frequently
76 rearrangements of chromosome 17 occur through meiosis, we screened an additional 48 progeny from the
77 same cross with a segment-specific PCR assay for ~500 bp regions of coding sequences at regular intervals
78 along chromosome 17, and found that anomalies of chromosome 17 were most likely restricted to the two
79 initially discovered progeny A66.2 and A2.2 (fig. 1A, supplementary fig. S1). A previous screen based on
80 restriction site associated DNA sequencing of a further 228 progeny from this cross revealed four progeny
81 with likely rearranged chromosome 17 (*14*). Finally, we assessed the frequency of rearranged chromosome
82 17 variants using whole-genome sequencing of 150 field-collected isolates matching the sampling location
83 of the parental isolates. Based on read coverage variation, the population carries likely three disomic or
84 partially duplicated variants of chromosome 17 (*15*).

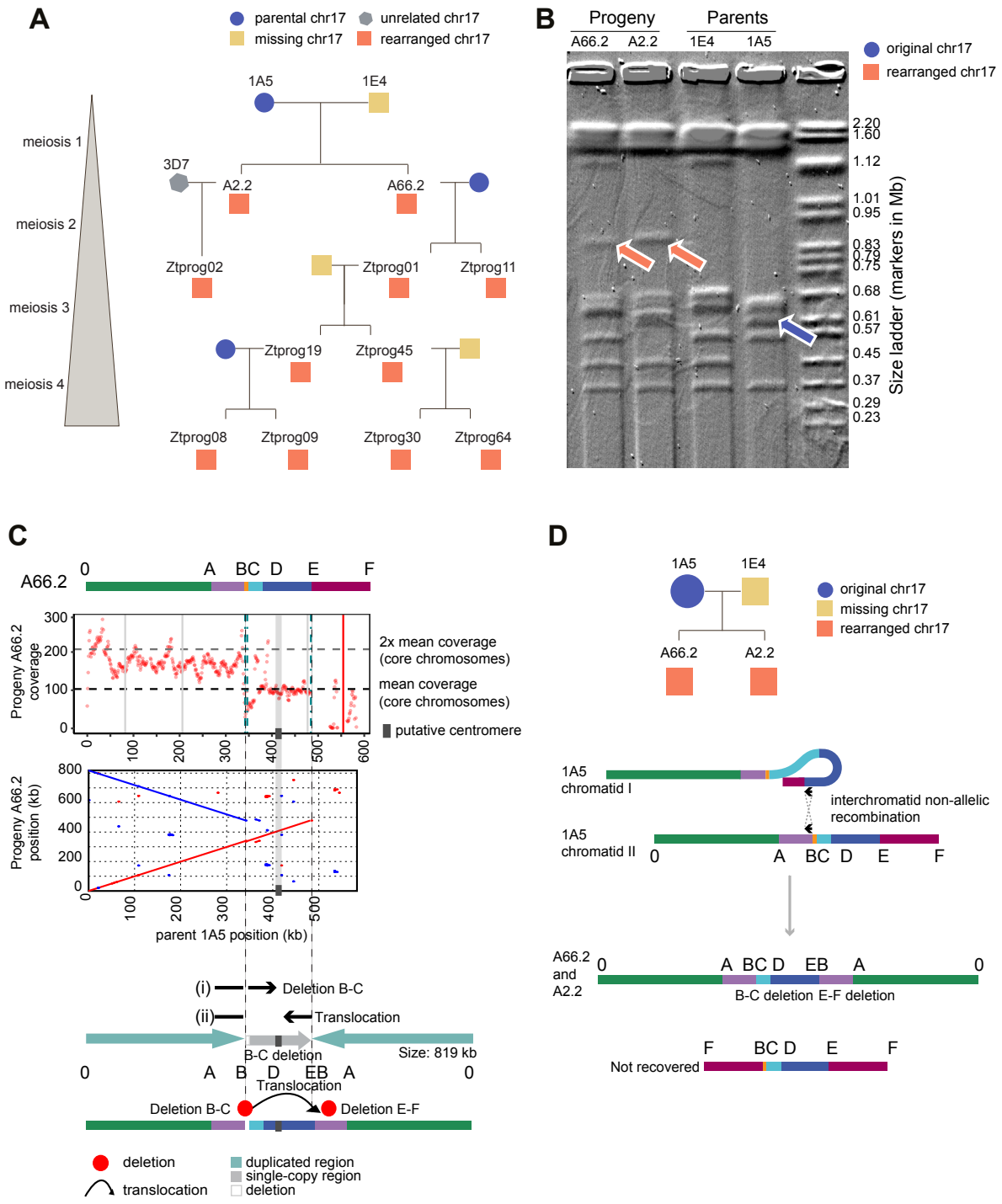


Figure 1: Pedigree analyses revealing repeated occurrence of enlarged chromosome 17. (A) Pedigree of four rounds of meiosis. The colors indicate whether the isolate carries a parental, unrelated (not carried by a parent), rearranged, or no chromosome 17 according to pulsed-field gel electrophoresis (PFGE) analysis. (B) PFGE of chromosomes from the progeny (A66.2 and A2.2) and the parents (1E4 and 1A5), adapted from (*7*). Orange arrows indicate the enlarged chromosome 17, the blue arrow indicates the parental chromosome 17. (C) The coverage and breakpoints of progeny A66.2 reads mapped to the parent 1A5. Black and grey horizontal dashed lines indicate mean

coverage and 2x mean coverage of the core chromosomes, respectively. Red dots indicate the mean coverage in 1 kb windows (regions with excessive, >300x coverage were removed). Dashed vertical lines indicate the chromosomal breakpoints identified from mapped reads (at positions B and E, turquoise and continuing in grey). Solid vertical lines indicate the positions of loci amplified by PCR (grey: amplification; red: no amplification). Dotplot of the assembled chromosome for progeny A66.2 compared to the parental chromosome. Inverted regions are indicated in blue. (i) and (ii) show the location and continuation of split reads. At the bottom is a schematic of the resulting enlarged chromosome in the progeny. (D) Schematic representation of the breakpoints and rearrangements between the two chromatids of 1A5 that generated the enlarged chromosome 17 (0–ABCDEBA–0) recovered in progeny A66.2. The hypothetical smaller rearrangement product (FBCDEF) was not observed.

To investigate at base-pair resolution how the enlarged chromosome 17 was generated in progeny A66.2 and A2.2, we sequenced the isolates with long-read PacBio technology. The recency of the duplication creates a significant challenge for the chromosomal assembly as duplicated regions are not expected to show sequence divergence and collapse during assembly. We used a combination of coverage and split long-read mapping approaches to overcome this challenge. First, we mapped reads from the progeny to the 1A5 genome and identified regions of chromosome 17 with higher than average coverage compared to the core chromosomes (1–13). Mean read coverage transitioned sharply along chromosome 17 (fig. 1C). We found that region 0–B in progeny A66.2 and A2.2 each had approximately double the mean coverage of the core chromosomes, implying that this region was duplicated (fig. 1C, supplementary fig. S2 and S3, supplementary table S1). Further supporting evidence included a breakpoint at the end of segment 0–B, with reads showing split mapping and linking two distinct regions of the chromosome (fig. 1C, supplementary fig. S2, supplementary table S2). Reads mapping to this position were evenly split between canonically mapping across positions B–C and split reads mapping from A–B and continue in an inverted orientation from position E towards D (fig. 1C, supplementary fig. S2). Taken together, we found that the region 0–B exists twice in the progeny genome and that the duplicated sequences are connected to two distinct locations of the chromosome at C and E (fig. 1C, supplementary fig. S2 and S3).

The rearranged chromosome 17 carries a single-copy region between C and E. A lack of coverage after position E towards F suggests that this region is absent in the rearranged chromosome. We confirmed the absence of region E–F by PCR (fig. 1B, supplementary fig. S2–3). A lack of coverage between B and C indicates a deletion of this region. Using the information on coverage and reads spanning distinct chromosomal regions, the chromosome 17 of progeny A66.2 and A2.2 are composed of region 0–B, followed by region C–E, and a second copy of region 0–B in an inverted orientation (fig. 1C; supplementary fig. S2, supplementary table S2). We identified the putative centromere to be located between position 405,779 and 415,898 based on sequence homology to the canonical chromosome 17 (*8*). Therefore, the resulting enlarged chromosome 17 is expected to carry a single copy of the centromeric region (fig. 1C). Both progeny most likely inherited an identical, enlarged chromosome 17 generated during the first round

130 of meiosis through the same non-allelic recombination between sister chromatids at locations B and E (fig
131 1D). The rearrangement should also have produced a second, shorter chromosomal variant. However, the
132 shorter variant was not recovered in progeny, despite being predicted to carry a centromere. We
133 reconstructed the sequence of the rearranged chromosome 17 to be 819 kb in length, matching the
134 approximate length identified by PFGE (fig. 1B).
135

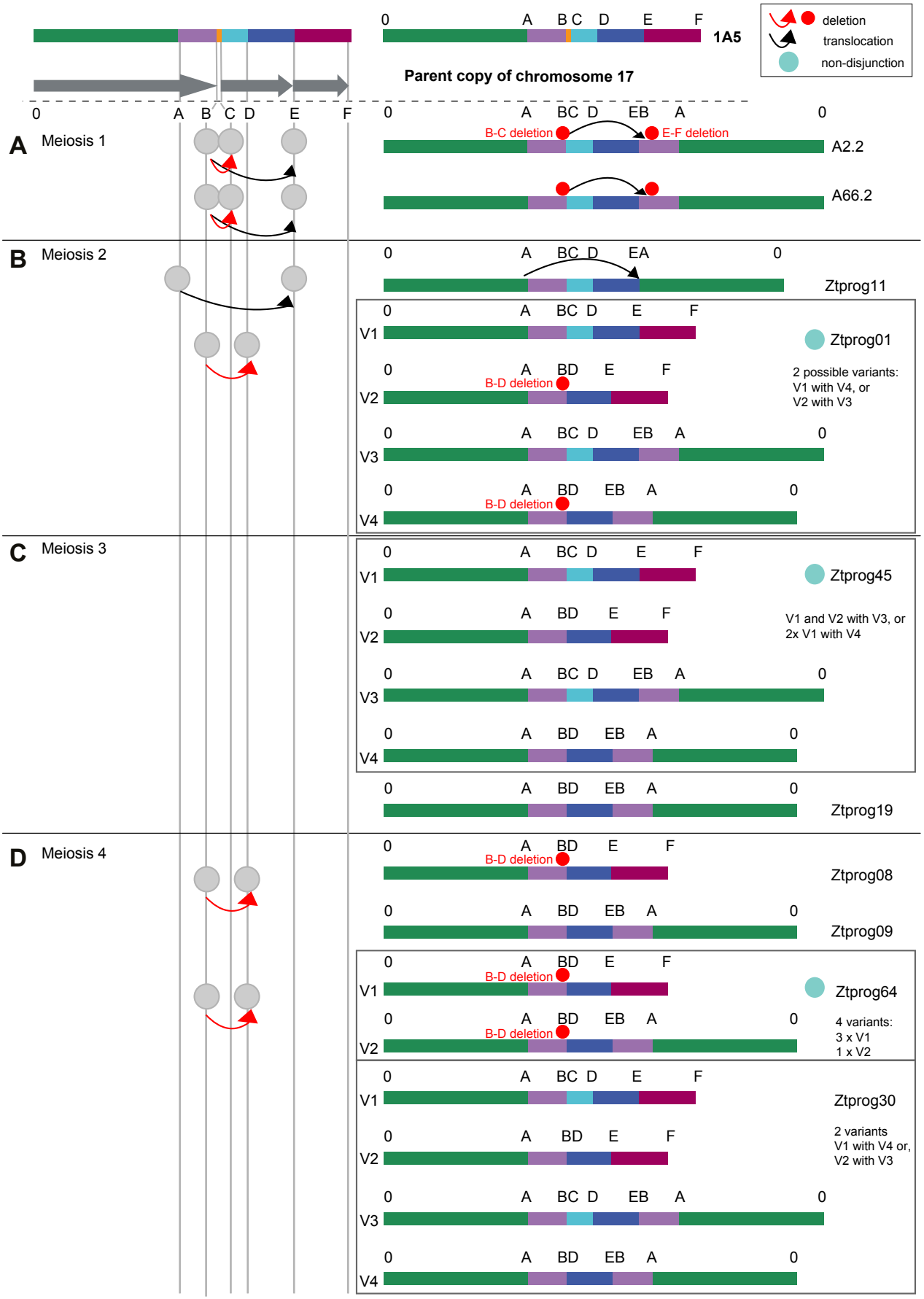

deletion
translocation
non-disjunction
1A5
Parent copy of chromosome 17
A
Meiosis 1
B-C deletion
E-F deletion
A2.2
A66.2
B
Meiosis 2
Ztprog11
Ztprog01
2 possible variants:
V1 with V4, or
V2 with V3
B-D deletion
C
Meiosis 3
Ztprog45
V1 and V2 with V3, or
2x V1 with V4
Ztprog19
D
Meiosis 4
Ztprog08
Ztprog09
Ztprog64
4 variants:
3 x V1
1 x V2
Ztprog30
2 variants
V1 with V4 or,
V2 with V3

Figure 2: Sequence of chromosome 17 rearrangements tracked through four rounds of meiosis. Chromosome rearrangements are shown according to breakpoint positions A–F identified through split read mapping. Arrows linking different letters A–F indicate translocations and deletions for the successive rounds of meiosis. Panels on the right show schematic representations of each chromosome 17 variant present in each progeny. If multiple variants were detected, variants are labelled V1–V4.

Sustained chromosome degeneration in subsequent rounds of meiosis

We investigated the fate of the enlarged chromosome 17 in further rounds of meiosis by performing several backcrosses (fig. 2A). New variants of chromosome 17 were already generated in the second round of meiosis (fig. 2B, 3B, supplementary fig. S3). Chromosome 17 of Ztprog11 carries a duplicated region, 0–A and the second copy of 0–A is joined to position E in an inverted orientation (fig. 2B, 3B, supplementary fig. S3, S4). Ztprog01 has two copies of chromosome 17, including a small and a large variant consistent with non-disjunction. One copy lacks the segment between B and D (fig 2B, fig 3B, supplementary fig. S3, S5). During the third round of meiosis, ZtProg19 stably inherited a chromosome 17 variant (fig. 2D, fig 3C, supplementary fig. S3, S7). Based on progeny and read mapping evidence, the second progeny ZtProg45 carries multiple variants. Variant V2 may have undergone non-disjunction (2x V1 is inherited with 1x V4) or one copy of V1 is present with one copy of V2 and V3 each, indicating rearrangements at breakpoints B and D as well as non-disjunction (fig. 2C, fig. 3C, supplementary fig. S3, S6). Therefore, two small variants are present together with one large variant.

In the last round of meiosis, the region B–D was deleted in Ztprog8 and Ztprog64 (fig. 2D, fig 3D and ii, supplementary fig. S3, S8, S9). Ztprog64 carries four chromosome 17 variants (three small and one large variant) that are missing the region B–D (fig. 2D, fig 3i and ii, supplementary fig. S3, S9). Ztprog9 most likely inherited the enlarged chromosome 17 without further rearrangements from the previous generation (fig. 2D, fig 3D, supplementary fig. S3, S10). Ztprog30 has the same chromosome complement as Ztprog01 (fig. 2D, fig 3E, supplementary fig. S3, S11). The enlarged chromosome 17 was highly unstable through further rounds of meiosis and degenerative cycles via non-allelic recombination as well as non-disjunction occurred. Rearrangements at all breakpoints except A were observed several times in the experiment.

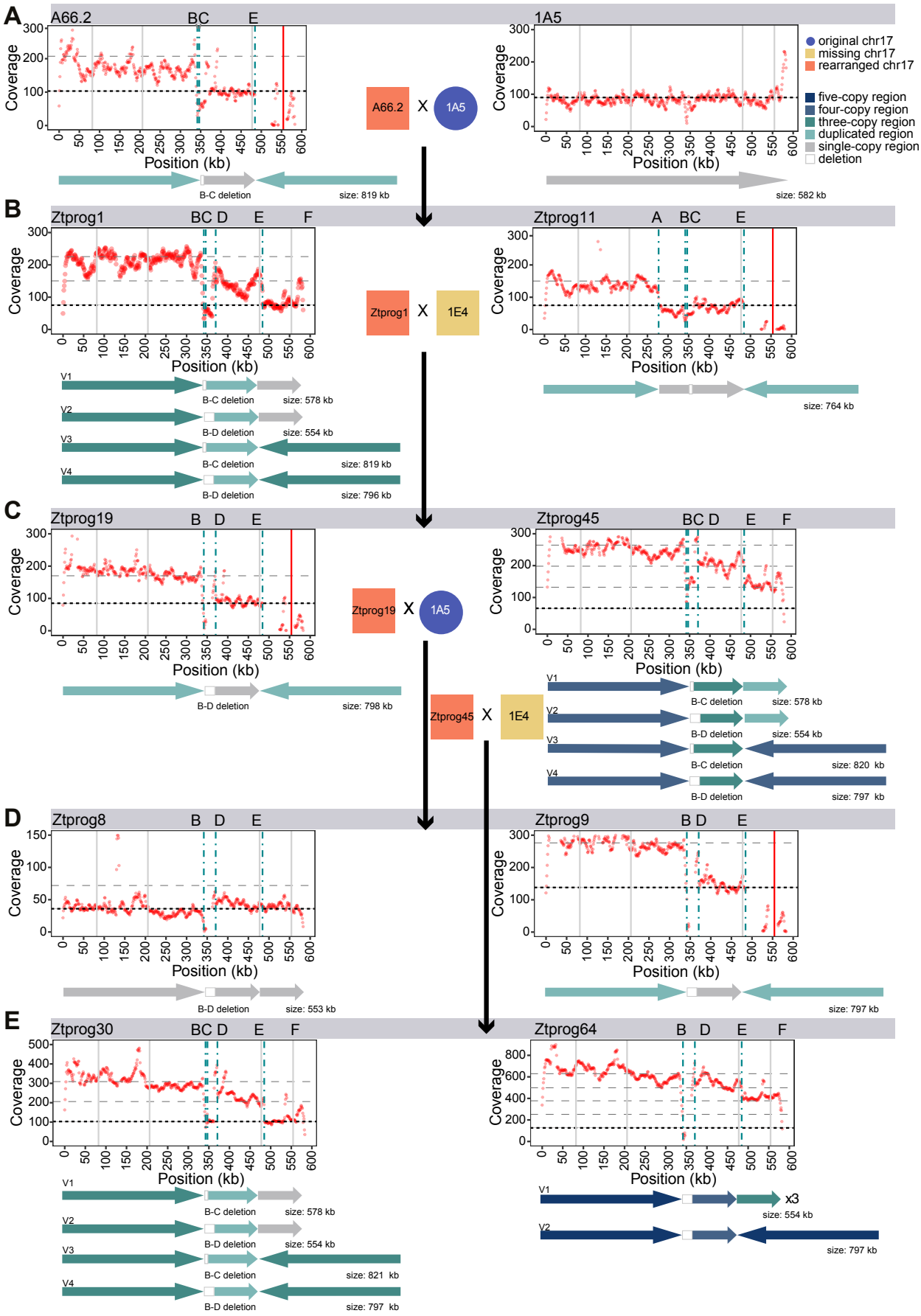
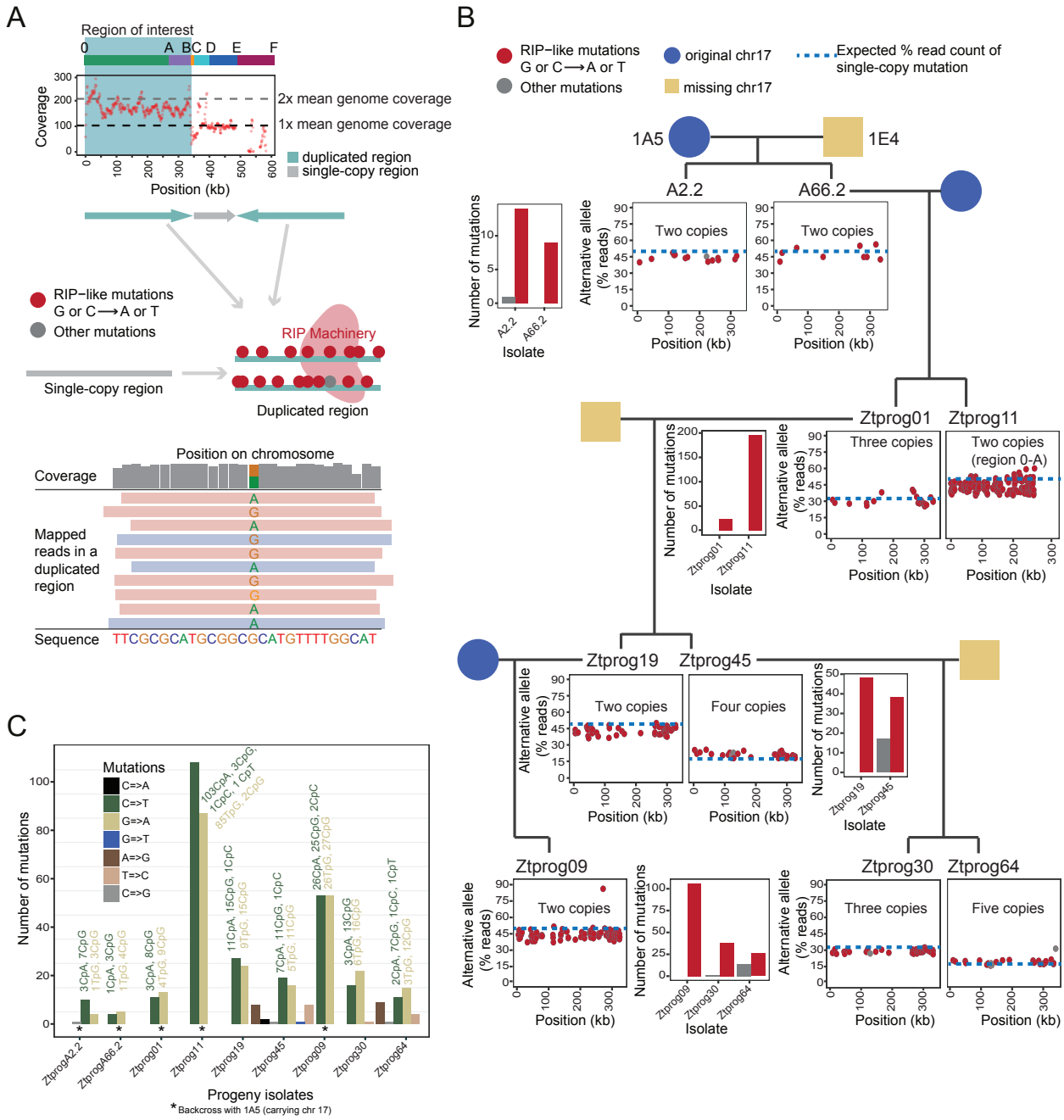

A
A66.2
BC
E
1A5
Coverage
Position (kb)
A66.2
X
1A5
original chr17
missing chr17
rearranged chr17
five-copy region
four-copy region
three-copy region
duplicated region
single-copy region
deletion
B-C deletion
size: 819 kb
size: 582 kb
B
Ztprog1
BC D
E
F
Ztprog11
A
BC
E
Ztprog1
X
1E4
V1
V2
V3
V4
B-C deletion
size: 578 kb
B-D deletion
size: 554 kb
size: 819 kb
size: 796 kb
size: 764 kb
C
Ztprog19
B
D
E
Ztprog45
BC D
E
F
Ztprog19
X
1A5
Ztprog45
X
1E4
B-D deletion
size: 798 kb
size: 820 kb
size: 797 kb
D
Ztprog8
B
D
E
Ztprog9
B D
E
size: 553 kb
size: 797 kb
E
Ztprog30
BC D
E
F
Ztprog64
B
D
E
F
x3
size: 821 kb
size: 797 kb

Figure 3: Reconstructed chromosome 17 variants based on sequence rearrangement breakpoints. For each progeny the mapped read coverage is shown relative to the 1A5 reference genome sequence. Black and grey horizontal dashed lines indicate mean coverage and 2x/3x/4x mean coverage of the core chromosomes, respectively. Red dots indicate the mean coverage in 1 kb windows. Vertical dashed lines indicate the chromosomal breakpoints A–F (see fig. 1 and 2). Solid vertical lines indicate the positions of loci amplified by PCR (grey: amplification; red: no amplification). Below each coverage plot, arrows show variants reconstructed for each progeny. If multiple variants were detected, variants are labelled with V1–V4. Variant labels are independent between progenies. Arrow color is based on the level of duplication (absent, single-copy or multiple-copy regions). The progeny is represented by (A) the progeny of the first meiotic round, (B) second, (C) third, and (D–E) fourth round.



176 **Figure 4: Characterization of repeat-induced point mutation (RIP) acting on duplicated sequences.** (A) Read
177 coverage along the chromosome 17 and evidence for a duplicated region 0–B (0–A in Ztprog11). The upper schematic
178 shows how RIP mutations are introduced into duplication regions, causing an excess of Adenine and Thymine with
179 RIP acting on Guanine or Cytosine. The lower schematic shows how mapped PacBio reads in duplicated regions were
180 used for SNP calling, revealing potentially RIP-like mutations. (B) Overview of progeny over generations with
181 barplots summarizing the number of RIP-like and other mutations detected per progeny generation in region 0–B (0–
182 A in Ztprog11). Dotplots show the percentage of reads carrying the alternative allele (parent 1A5). The identified copy
183 numbers of the region 0–B (0–A in Ztprog11) is indicated. Blue dashed lines show the expected percentage of reads,
184 confirming a new mutation carried by at least one copy of the duplicated regions. (C) Breakdown of all mutations
185 detected in duplicated regions of each progeny. Mutations at particular dinucleotides are summarized individually per
186 progeny.
187

188 ***Degeneration of the duplicated sequence through RIP***

189 Fungal genomes encode repeat-induced point mutations (RIP), a defense mechanism targeting duplicated
190 sequences that dramatically reduces TE activity. We predicted that the presence of a large, duplicated region
191 in chromosome 17 would trigger RIP (fig. 4A). To detect RIP-like mutations, we mapped reads from
192 progeny to the parental chromosome 17 and identified regions that with transitions from G/C to A/T (fig.
193 4B; supplementary fig. S4). We found an overrepresentation of RIP-like mutations in region 0–B. As a
194 control, we analyzed the rearrangement-free progeny Ztprog08 for RIP-like mutations and found indeed no
195 overrepresentation of RIP-like mutations, indicating RIP not being active. Hence, the genomic defenses are
196 specifically active against the new duplications appearing in degenerate chromosome 17 variants. RIP-like
197 mutations predominantly targeted CpA dinucleotides (or TpG) (fig. 3C), matching evidence from other
198 model fungi (*16*). Additionally, CpG dinucleotides were also frequently targeted. New RIP-like mutations
199 were detected after a single round of meiosis, suggesting that the RIP genome defense did not yet reach
200 saturation. RIP-like mutations never targeted the putative centromere (405,779–415,898 bp; supplementary
201 fig. S4). Our results show that massive rearrangements and non-disjunction triggers the genome defense
202 RIP, contributing to an increased mutation rate in the duplicated region.

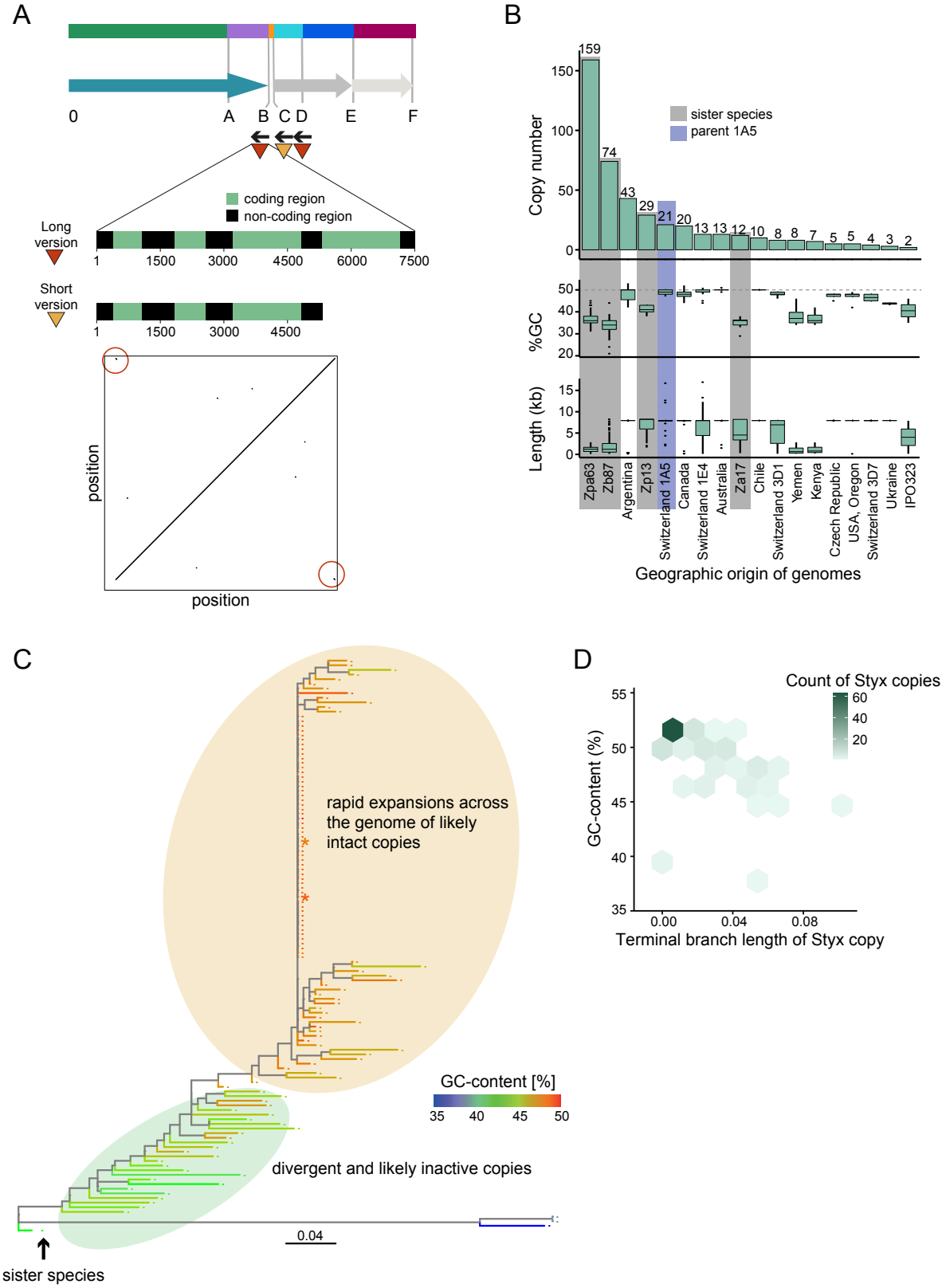

A
0
A
B
C
D
E
F
coding region
non-coding region
Long version
1
1500
3000
4500
6000
7500
Short version
1
1500
3000
4500
position
position
B
Copy number
159
74
43
29
21
20
13
13
12
10
8
8
7
5
5
4
3
2
sister species
parent 1A5
%GC
Length (kb)
Zpa63
Zb87
Argentina
Zp13
Switzerland 1A5
Canada
Switzerland 1E4
Australia
Za17
Chile
Switzerland 3D1
Yemen
Kenya
Czech Republic
USA, Oregon
Switzerland 3D7
Ukraine
IPO323
Geographic origin of genomes
C
rapid expansions across the genome of likely intact copies
GC-content [%]
35
40
45
50
divergent and likely inactive copies
0.04
sister species
D
Count of Styx copies
60
40
20
GC-content (%)
55
50
45
40
35
0.00
0.04
0.08
Terminal branch length of Styx copy

Figure 5: Recent expansion of the *Styx* transposable element underpinning the chromosomal rearrangements. (A) The length and location of coding regions of the long and short copies of *Styx*. Dotplot of the consensus sequence showing duplicated regions (red circle). (B) Copy number, GC-content, and length distribution of *Styx* copies recovered from a global panel of *Z. tritici* reference genomes. A dark grey background indicates genomes of sister species *Z. ardabiliae* (Za17), *Z. passerini* (Zpa63), *Z. brevis* (Zb87), and *Z. pseudotritici* (Zp13). (C) Phylogeny of *Styx* copies in *Z. tritici*, with *Z. pseudotritici* as an outgroup. The color scale indicates the GC content. The two full-length copies found on chromosome 17 of the parent 1A5 are indicated by a star. (D) Density plot of branch length against GC-content of individual *Styx* copies. The most recent burst of *Styx* copies show small branch lengths and high GC-content.

Chromosomal rearrangements triggered by a recent TE burst

To understand the mechanism triggering the initial chromosome 17 rearrangement and sustaining the degeneration, we examined the breakpoint sequences. We identified a full copy of the DNA transposon *Styx* (also known as REP9) at position B and a partial copy of the same family at position C in the parental chromosome of 1A5 (fig. 5A). *Styx* was previously described as a negative regulator of virulence (*17*). Furthermore, *Styx* was shown to proliferate in the progeny and pathogen populations across continents (*18*, *19*). At position B, the element is at 582 bp from the breakpoint. A second complete copy of *Styx* was found at position D (~8 kb in length) (fig. 5A). *Styx* has 21 copies in the 1A5 genome, including three copies on chromosome 17 (fig. 5B). Using ncbi blast, we predicted four putative coding sequences in full-length copies of *Styx*, of which one shows weak homology to RNase H and an integrase (fig. 5A) (*18*). We analyzed *Styx* copy numbers in 19 completely assembled genomes of *Z. tritici* (*13*) and in the genomes of the sister species *Z. passerinii*, *Z. brevis* and *Z. pseudotritici* (*20*) . *Styx* is present in higher copy numbers in the sister species but nearly absent in genomes sampled from the center of origin of *Z. tritici* (*18*). Our findings support a scenario of *Styx* being present in the common ancestor at low copy numbers and undergoing independent bursts in the sister species, as well as in *Z. tritici* populations in North and South America, Europe, and Australia, resulting in increased copy numbers in the parental and other isolates (fig. 5B) (*18*). The TE copy on chromosome 17 was introduced following a recent burst (fig. 5C). TE copies created during the burst are characterized by short terminal branch lengths and high GC-content (fig. 5D). The creation of the enlarged chromosome 17 was likely mediated by non-allelic homologous recombination between *Styx* copies at positions B and a different sequence with microhomology at position E. We examined the regions 1 kb up- and downstream from position A, B, and E for similar repetitive sequences. Positions B and E carry a similar 6-bp repeat that may have ultimately triggered the rearrangement (supplementary files S1–2). Finally, near position E the repeat is 34 bp away from the breakpoint. Similarly, positions A and E had two similar 3-bp repeats (supplementary file S2 and S3).

Discussion

240 We reveal the dynamics of an unusual chromosomal rearrangement in a fungal plant pathogen. Using split
241 long-reads, we identified the exact breakpoints and retraced the degeneration through four rounds of
242 meiosis. We propose that the primary degenerative rearrangement was caused by non-allelic recombination
243 between a TE and a region with microhomology to the TE. The degenerated chromosome was composed
244 of a large, duplicated, and inverted region connected to a single-copy chromosomal segment near the
245 centromere. The sequences serving as triggers for the rearrangement in the first round of meiosis was likely
246 the transposable element *Styx*, being present in multiple copies on chromosome 17 and regions and sharing
247 microhomology. As an accessory chromosome without known function, chromosome 17 is likely under
248 relaxed selection, and chromosomal rearrangements are tolerated. However, chromosomal rearrangements
249 co-locating with *Styx* copies were also detected in core chromosomes of the same cross (*18*). Subsequent
250 rounds of meiosis increased the spectrum of rearrangement breakpoints consistent with runaway
251 chromosomal degeneration. Chromosomal segments affected by the rearrangement were repeatedly
252 affected multiple times and three of these segments co-localize with copies of *Styx*. Concurrent with the
253 degeneration, non-disjunction events of chromosome 17 increased along the progeny, characterized by
254 disomy and trisomy. The frequency of non-disjunction events increased with the presence of chromosome
255 17 variants in both parents, suggesting that opportunities for mispairing of chromatids increases the
256 likelihood of aberrant segregation. The degenerating chromosome was furthermore affected by the genomic
257 defense mechanism RIP, introducing mutations into recently duplicated sequences. RIP on very large
258 duplicated sequences might have a titration effect, leading to less dense RIP mutations in smaller duplicated
259 sequences (*21*). Overall, the identity of the paired parental genotypes had a major influence on the likelihood
260 of rearrangements observed in the progeny.
261
262 The copy-number amplification in progeny Ztprog64 and Ztprog30 shares hallmarks of breakage-fusion-
263 bridge (BFB) cycles observed in many cancer lines (*22–27*). Following the initial degeneration,
264 chromosome 17 was stabilized in pairings of a single rearranged chromosome variant with a parental
265 chromosomal variant (*i.e.,* Ztprog8 and Ztprog9). This stabilization suggests that the progression of the
266 degenerative cycle was interrupted through proper segregation. The chromosome 17 variant in Ztprog8 was
267 generated through non-allelic homologous recombination between two *Styx* copies. The repair of
268 chromosomes through recombination between repeats is known to produce intermediary chromosomes that
269 can ultimately produce stable karyotypes (*28*). The reconstructed chromosome 17 variants pinpointed the
270 most likely sequence triggers for the observed rearrangements. The observation that specific locations on
271 the chromosome were repeatedly involved in creating new rearrangements during chromosome
272 degeneration indicates that these locations are fragile sites (*i.e.* sites frequently co-locating with
273 chromosome rearrangements). Three out of four of these fragile sites co-locate with copies of *Styx*. The

tendency of *Styx* to trigger chromosomal degeneration is likely widespread within the species, as the same TE is also responsible for an inter-chromosomal rearrangement (*18*). Isolates collected near the geographic center of origin of the pathogen in the Middle East carry no or only few *Styx* copies in the genome and the expansion to higher copy-numbers appears restricted to European genotypes and their descendants (*13*). A separate and likely independent burst of amplification of *Styx* appears to have occurred in the sister species of *Z. tritici*, where no chromosomal rearrangements on chromosome 17 have been described yet (*18*, *29*). The significant consequences for chromosomal integrity could mean that the activity of *Styx* and the presence of fragile sites are under strong selection.

Taken together, our results indicate that specific TE sequences can trigger runaway chromosome degeneration. Non-allelic homologous recombination drives the deleterious rearrangements at the onset of the process with non-disjunction events in subsequent rounds of meiosis. Specific regions on chromosomal segments are preferentially amplified, consistent with patterns observed during degenerative BFB cycles in cancer cell lines (*24*, *30*). BFB was first discovered in maize by McClintock in dicentric chromosomes going through cycles of degeneration (*31*, *32*) and are also known to occur in animals (*33*, *34*) and fungi (*35*, *36*). BFB cycles are initiated via telomere-telomere fusions of chromosomes with degraded or missing telomeres (*37*). The centromeres of the dicentric chromosome are pulled in opposite directions during anaphase, generating a bridge that breaks apart and results in daughter cells with different lengths of the chromosome that lack telomeres (*25*, *33*, *38–40*). BFB cycles tend to be self-perpetuating through repeated fusion and breakage events. At the karyotype level, chromosome 17 undergoes cycles consistent with BFB cycles. However, the likely absence of centromere duplications is inconsistent with classic BFB cycles. We pinpoint an alternative mechanism driving amplification of chromosomal regions in a pattern resembling BFB and involving ectopic recombination. Recombination serves both as the initial trigger to create unstable chromosomes and as the mechanism to maintain the degenerative process. Non-disjunction amplifies the process following the initial meiosis, producing the aberration. We show that large-scale chromosome rearrangements can spontaneously occur in natural pairings of fungal individuals, bearing hallmarks of degenerative processes in somatic cell lines such as those observed in cancers.

305 ## Materials and Methods

306 ***Establishment of a four-generation progeny***

307 To detect how chromosomal rearrangements evolved on chromosome 17, we analyzed four generations of
308 crosses, including two backcrosses of the haploid progeny A66.2 (fig. 1A). The initial cross between
309 parental strains 1E4 and 1A5, and subsequent crosses were described earlier (*7*, *41*). We used previously
310 generated telomere-to-telomere PacBio assembly of parental strains 1E4, 1A5 and 3D7 (*9*, *11*). Parental
311 genome assemblies were validated using high-density genetic maps (*7*, *42*). Crosses were performed by
312 coinfecting wheat leaves with asexual conidia from the parental strains of opposite mating types, according
313 to an established crossing protocol (*43*): Conidia of both parents were sprayed onto wheat plants in equal
314 concentration and incubated outdoors for 40–60 days. Ascospores were isolated over several days by
315 incubating infected wheat leaves on wet filter paper inside Petri dishes. Wheat leaves were covered with
316 upside down Petri dish lids filled with water agar, enabling the capture of vertically ejected ascospores.
317 Ascospores captured on the water agar were left to germinate and inspected for contaminants. Only progeny
318 isolates from single ascospores were selected. Each germinating ascospore was transferred to an individual
319 culture plate for clonal propagation. The mycelium produced by each ascospore was used for DNA
320 extraction. Progeny mycelium was grown in YSB (yeast sucrose broth) liquid medium for 6–7 days at 20°C
321 prior to DNA extraction.

322

323 ***Chromosome segment PCR assay***

324 In order to assess the presence-absence polymorphism of chromosome 17 segments, we used previously
325 designed PCR assays to amplify ~500 bp regions of coding sequences at regular intervals along the
326 chromosome 17 of reference strain IPO323 (*7*). Detailed information on primer binding sites in the genome
327 of 1A5 is available in supplementary table S3. PCR reactions were performed in 20 µl volumes with 5–10
328 ng genomic DNA, 0.5 mM of each primer, 0.25 mM dNTP, 0.6 U Taq polymerase (DreamTaq, Thermo
329 Fisher, Inc.), and the corresponding PCR buffer. In order to avoid false negatives, we included a primer
330 pair for a conserved microsatellite locus in each PCR mix (*44*). The amplification protocol was described
331 in more detail previously (*7*). Successful PCRs produced an additional band that was clearly distinguishable
332 from the PCR product associated with the amplified chromosome region. PCR products were analyzed on
333 agarose gels. Data was visualized using the *R* package *gplots* (https://github.com/talgalili/gplots).

334

335 ***DNA extraction for PacBio sequencing***

336 Progeny DNA from each cross was extracted using a modified version of the cetyltrimethylammonium
337 bromide (CTAB) DNA extraction protocol developed for plant DNA extractions (*45*). Fungal cultures were
338 grown for 5–7 days in YSB broth and lyophilized overnight. Approximately 60–100 mg of dried material

339 was crushed with a mortar and pestle. The phenol-chloroform-isoamyl alcohol extraction step was
340 performed twice and the washing step three times. In the last step, the DNA pellet was resuspended in
341 100 µl of sterile water.
342
343 ***Preparation of fungal material for molecular karyotyping***
344 DNA from intact chromosomes was extracted from conidia from fungal cultures, embedded in agarose gels
345 by the *in situ* digestion of cell walls, using a modified non-protoplasting method (*46*). We included seven
346 *Z. tritici* isolates confirmed with PCR testing to have inherited chromosome 17 from each of the crosses.
347 Isolates were transferred from stocks maintained in glycerol at -80°C to yeast malt agar plates and incubated
348 for 3–4 days in the dark at 18°C. Conidia were then isolated by washing the plates with sterile water and
349 transferring 600–800 µl of suspended conidia to new YMA plates. The plates were again incubated for 2 to
350 3 days as described above. Conidia were harvested by washing the plates with sterile distilled water and
351 filtered through sterile Miracloth (Calbiochem, La Jolla CA, USA) into 50 ml Falcon tubes. The volume
352 was adjusted to 50 ml by adding more distilled water and the suspension was centrifuged at 3750 rpm at
353 room temperature for 15 min with a clinical centrifuge (Allegra X-12R, Beckman Coulter, Brea CA, USA).
354 The pellets were resuspended in 1–3 ml TE buffer (10 mM Tris-HCL, pH 7.5; 1 mM EDTA, pH 8.0) and
355 vortexed gently. The spore concentration of the solution was calculated using a Thoma haematocytometer
356 cell counter. The 1.5 ml spore suspensions with a concentration between 8×10^7 to 2×10^8 spores/ml were
357 transferred to 50 ml Falcon tubes and incubated at 55°C in a water bath for a few minutes. We added 1.5
358 ml pre-warmed (55°C) low-melting-point agarose prepared in TE Buffer (2% w/v; molecular biology grade,
359 Biofinex, Switzerland). The solution was mixed by gentle pipetting. An aliquot of 500 µl was solidified on
360 ice for approximately 10 min in a pre-cooled plug casting mold (BioRad Laboratories, Switzerland).
361 Agarose plugs were incubated in 15 ml Falcon tubes containing 5 ml of a lysing solution (0.25 M EDTA,
362 pH 8.0, 1.5 mg/mL protease XIV (Sigma, St. Louis MO, USA), 1.0% sodium dodecyl sulfate (Fluka,
363 Switzerland)). Plugs were incubated for ~24h at 55°C. The lysing solution was changed once after ~18h
364 and gently mixed every few hours. Plugs with whole chromosomal DNA were washed three times for 15–
365 20 min in ~5ml of a 0.1 M EDTA (pH 9.0) solution and then stored in the same solution at 4°C until use.
366
367 ***Pulsed-field gel electrophoresis***
368 To detect length differences in chromosome 17, pulsed-field gel electrophoresis (PFGE) was performed
369 using a BioRad CHEF II apparatus (BioRad Laboratories, Hercules CA, USA). Chromosomal plugs were
370 placed in the wells of a 1.2% (w/v) agarose gel (Invitrogen, Switzerland) to separate small chromosomes
371 up to 1Mb. Chromosomes were separated at 13°C in 0.56× Tris-borate-EDTA Buffer (*47*) at 200 V with a

372 60–120 s pulse time gradient for 24–26 h. Gels were stained in ethidium bromide (0.5 mg/ml) for 30 min.
373 De-staining was performed in water for 5–10 min. Photographs were taken under ultraviolet light with a
374 Molecular Imager (Gel Doc XR+, BioRad, Switzerland). As size standards, we used chromosome
375 preparations of *Saccharomyces cerevisiae* (BioRad, Switzerland).
376
377 ***Southern hybridization for A66.2 and A2.2***
378 Southern blotting and hybridization were performed following standard protocols (*47*). First, hydrolysis
379 was performed in 0.25 M HCl for 30 min, then DNA was transferred onto Amersham Hybond-N+
380 membranes (GE Healthcare, Switzerland) overnight under alkaline conditions. DNA was heat-fixed onto
381 the membranes at 80°C for 2 h. Membranes were prehybridized overnight with 25 ml of a buffer containing
382 20% (w/v) SDS, 10% BSA, 0.5 M EDTA (pH 8.0), 1 M sodium phosphate (pH 7.2) and 0.5 ml of sonicated
383 fish sperm solution (Roche Diagnostics, Switzerland). Probes were labeled with ^{32}P by nick translation
384 (New England Biolabs, Inc.) following the manufacturer's instructions. Hybridization was performed
385 overnight at 65°C. Blots were subjected to stringent wash conditions with a first wash in 16 X SSC and
386 0.1% SDS and a second wash with 0.26 X SSC and 0.1% SDS. Both washes were performed at 60°C.
387 Membranes were exposed to X-ray film (Kodak BioMax MS) for 2 to 3 days at -80°C. We used the same
388 probe as described earlier (supplementary table S4) (*7*).
389
390 ***PacBio library preparation***
391 PacBio SMRTbell libraries were prepared using 15–31 µg of high-molecular-weight DNA. The libraries
392 were size-selected with an 8 kb cutoff on a BluePippin system (Sage Science, Inc.). After selection, the
393 average fragment length was 15 kb. PacBio sequencing was run on a PacBio RS II instrument or Sequel at
394 the Functional Genomics Center, Zurich, Switzerland using P4/C2 and P6/C4 chemistry, respectively.
395
396 ***Assembly of chromosome 17 and breakpoint-junction analyses***
397 Each chromosome 17 for all four generations was then assembled (fig. 1A). We mapped the reads to the
398 reference genome 1A5 using minimap2 version 2.17 (*48*) with the parameters --secondary=no -ax map-pb.
399 We compared the coverage of regions of chromosome 17 to the mean coverage of the core chromosomes
400 (1–13) to estimate copy numbers of distinct regions. We then identified breakpoints by analyzing regions
401 with >15 reads that either end or start at a specific position using BEDTools bamtobed version 2.29.2 (*49*)
402 and extracted split reads in this region. We then assembled draft chromosomes by using the structural
403 information from reads showing split alignments, and by joining individual breakpoints (supplementary
404 table S2). Reads were mapped to the assembled chromosomes with minimap2, and we counted the number

405 of reads spanning each established junction point (supplementary table S2). Established chromosome 17
406 assemblies were error-corrected with Quiver version 2.1.2 (*50*).
407
408 ***Characterization of sequence features of breakpoints***
409 Transposable elements consensus sequences were annotated on the parent 1A5 chromosome 17 sequence
410 with RepeatMasker version 2.10.0 based on previously described TE family consensus sequences (*13*).
411 The cut-off value was set to 250 and both simple repeats and low-complexity regions were discarded. We
412 used nucmer version 4.0.0rc1 and dotPlotly (https://github.com/tpoorten/dotPlotly/) to visualize the *Styx*
413 TE consensus sequence (*51*). To further classify *Styx*, we analyzed conserved domains in the consensus
414 sequence with online BLASTx on the nonredundant NCBI protein database. We detected four putative
415 coding regions consistent with previous analyses of the TE (*18*). As a proxy for repeat-induced point
416 mutations, we calculated the GC-content for each copy with *geecee* from EMBOSS version 6.6.0 (*52*). To
417 search for microhomology at rearrangement breakpoints, we analyzed tandem repeats within 1000 bp of
418 the breakpoint locations A–E using *mreps* with the parameters -exp 3 and -res 5 to allow for fuzzy detection
419 of degenerate repeats (*53*).
420
421 ***Genome-wide description of the Styx TE family***
422 We extracted all annotated copies of *Styx* in the 19 reference-quality *Z. tritici* genomes and the sister species
423 (*13*, *20*) with SAMtools faidx version 1.9 (*54*). Multiple sequence alignments of all copies were made with
424 MAFFT version 7.453 using the following parameters: --reorder --localpair --maxiterate 1000 --
425 nomemsave --leavegappyregion (*55*). We located the putative coding region 4 in the multiple sequence
426 alignment and extracted the sequence with extractalign from EMBOSS. We excluded empty hits (*i.e.* not
427 containing coding regions) with trimAl version 1.4rev15 and sequences with more than 20% of gap sites
428 with SeqKit version 0.11.0 (*56*, *57*). To remove large blocks of gaps and regions that represent rare
429 insertions, we used Gblocks version 0.91b with the parameters: -t=d -b4=5 -b5=h (*58*). We used RAxML
430 version 8.2.12 to create phylogenetic trees in three rounds (*7*): First, we generated 20 ML trees each with a
431 different starting tree and extracted the tree with best likelihood with the following parameters: raxmlHPC-
432 PTHREADS-SSE3 -T 4 -m GTRGAMMA -p 12345 -# 10 --print-identical-sequences. Second, we made a
433 bootstrap search for support values with the following parameters: raxmlHPC-PTHREADS-SSE3 -T 4 -m
434 GTRGAMMA -p 12345 -b 12345 -# 50 --print-identical-sequences. Finally, we drew bipartitions on the
435 best ML tree with bootstrapping: raxmlHPC-PTHREADS-SSE3 -T 4 -m GTRGAMMA -p 12345 -f b --
436 print-identical-sequences. We imported the tree into R with read.tree from the package treeio version 1.10.0
437 (https://github.com/YuLab-SMU/treeio), converted it to a tibble object with as.tibble from package tibble
438 version 3.0.1 in tidyverse version 1.3.0 and added the GC-content per sequence as a variable with left_join

from the package *dplyr* version 0.8.5 in tidyverse. The tree was then recreated with as.phylo and as.treedata in *treeio* (*59*, *60*). We chose a sequence from *Z. pseudotritici* as an outgroup to root the *Z. tritici* tree. We visualized the tree with GGTree version 2.0.1 (*61*, *62*).

Repeat-induced point mutation analysis

To detect the impact of RIP in the progeny, we checked for biases of mutations for specific dinucleotides. We mapped PacBio reads to the reference genome 1A5 using minimap2 (*48*) as described above. We performed SNP calling with the software Longshot version 0.4.0 and included a minimum mapping quality cutoff of -q 30 (*63*). All SNPs with a "dn" tag, indicating mapping issues, were removed. We allowed for 20% errors in the uncorrected PacBio reads. We expected the alternate allele at polymorphic sites to be shared by ≥80% of the mapped reads for single-copy regions, shared by ≥40% of the reads in regions with evidence for duplications (based on read coverage), shared by ≥27%, ≥20% and ≥16% of the reads in regions with three, four and five copies (based on read coverage), respectively. We also excluded regions with ≤50% or ≥150% of the expected read coverage for further analyses to reduce erroneous variant calls due to inconsistent read mapping.

References

1. T. Hassold, P. Hunt, To err (meiotically) is human: the genesis of human aneuploidy. *Nat Rev Genet* **2**, 280–291 (2001).

2. S. Negrini, V. G. Gorgoulis, T. D. Halazonetis, Genomic instability — an evolving hallmark of cancer. *Nat Rev Mol Cell Biol* **11**, 220–228 (2010).

3. S. G. Durkin, T. W. Glover, Chromosome Fragile Sites. *Annu. Rev. Genet.* **41**, 169–192 (2007).

4. R. Kumar, G. Nagpal, V. Kumar, S. S. Usmani, P. Agrawal, G. P. S. Raghava, HumCFS: a database of fragile sites in human chromosomes. *BMC Genomics* **19**, 985 (2019).

5. J. P. M. Camacho, T. F. Sharbel, L. W. Beukeboom, B-chromosome evolution. *Phil. Trans. R. Soc. Lond. B* **355**, 163–178 (2000).

6. T. Langner, A. Harant, L. B. Gomez-Luciano, R. K. Shrestha, A. Malmgren, S. M. Latorre, H. A. Burbano, J. Win, S. Kamoun, Genomic rearrangements generate hypervariable mini-chromosomes in host-specific isolates of the blast fungus. *PLoS Genet* **17**, e1009386 (2021).

7. D. Croll, M. Zala, B. A. McDonald, Breakage-fusion-bridge cycles and large insertions contribute to the rapid evolution of accessory chromosomes in a fungal pathogen. *PLOS Genet* **9**, e1003567 (2013).

8. K. Schotanus, J. L. Soyer, L. R. Connolly, J. Grandaubert, P. Happel, K. M. Smith, M. Freitag, E. H. Stukenbrock, Histone modifications rather than the novel regional centromeres of *Zymoseptoria tritici* distinguish core and accessory chromosomes. *Epigenetics Chromatin* **8**, 41 (2015).

476 9. C. Plissonneau, A. Stürchler, D. Croll, The evolution of orphan regions in genomes of a fungal pathogen of
477 wheat. *mBio* **7**, e01231-16 (2016).

478 10. S. Fouché, C. Plissonneau, D. Croll, The birth and death of effectors in rapidly evolving filamentous pathogen
479 genomes. *Curr. Opin. Microbiol.* **46**, 34–42 (2018).

480 11. C. Plissonneau, F. E. Hartmann, D. Croll, Pangenome analyses of the wheat pathogen *Zymoseptoria tritici*
481 reveal the structural basis of a highly plastic eukaryotic genome. *BMC Biology* **16**, 5 (2018).

482 12. A. Sánchez-Vallet, S. Fouché, I. Fudal, F. E. Hartmann, J. L. Soyer, A. Tellier, D. Croll, The genome biology
483 of effector gene evolution in filamentous plant pathogens. *Annual Review of Phytopathology* **56**, 21–40
484 (2018).

485 13. T. Badet, U. Oggenfuss, L. Abraham, B. A. McDonald, D. Croll, A 19-isolate reference-quality global
486 pangenome for the fungal wheat pathogen *Zymoseptoria tritici*. *BMC Biology* **18**, 12 (2020).

487 14. S. Fouché, C. Plissonneau, B. A. McDonald, D. Croll, Meiosis Leads to Pervasive Copy-Number Variation
488 and Distorted Inheritance of Accessory Chromosomes of the Wheat Pathogen Zymoseptoria tritici. *Genome*
489 *Biol Evol* **10**, 1416–1429 (2018).

490 15. N. K. Singh, P. Karisto, D. Croll, Population-level deep sequencing reveals the interplay of clonal and sexual
491 reproduction in the fungal wheat pathogen Zymoseptoria tritici. *Microbial Genomics* **7** (2021).

492 16. E. Gladyshev, "Repeat-Induced Point Mutation and Other Genome Defense Mechanisms in Fungi" in *The*
493 *Fungal Kingdom*, J. Heitman, B. J. Howlett, P. W. Crous, E. H. Stukenbrock, T. Y. James, N. A. R. Gow, Eds.
494 (ASM Press, Washington, DC, USA, 2017; http://doi.wiley.com/10.1128/9781555819583.ch33), pp. 687–
495 699.

496 17. C. Wang, A. W. Milgate, P. S. Solomon, M. C. McDonald, The identification of a transposon affecting the
497 asexual reproduction of the wheat pathogen *Zymoseptoria tritici*. *Mol Plant Pathol* **22**, 800–816 (2021).

498 18. T. Badet, S. M. Tralamazza, A. Feurtey, D. Croll, Recent reactivation of a pathogenicity-associated
499 transposable element is associated with major chromosomal rearrangements in a fungal wheat pathogen.
500 *Nucleic Acids Research* **52**, 1226–1242 (2024).

501 19. A. Feurtey, C. Lorrain, M. C. McDonald, A. Milgate, P. S. Solomon, R. Warren, G. Puccetti, G. Scalliet, S. F.
502 F. Torriani, L. Gout, T. C. Marcel, F. Suffert, J. Alassimone, A. Lipzen, Y. Yoshinaga, C. Daum, K. Barry, I.
503 V. Grigoriev, S. B. Goodwin, A. Genissel, M. F. Seidl, E. H. Stukenbrock, M.-H. Lebrun, G. H. J. Kema, B.
504 A. McDonald, D. Croll, A thousand-genome panel retraces the global spread and adaptation of a major fungal
505 crop pathogen. *Nat Commun* **14**, 1059 (2023).

506 20. A. Feurtey, Lorrain, Cécile, Croll, Daniel, Eschenbrenner, Christoph, Freitag, Michael, Habig, Michael,
507 Haueisen, Janine, Möller, Mareike, Schotanus, Klaas, Stukenbrock, Eva H, Genome compartmentalization
508 predates species divergence in the plant pathogen genus Zymoseptoria. *BMC Genomics* **21**, 588 (2020).

509 21. P. K. Singh, D. P. Kasbekar, Titration of repeat-induced point mutation (RIP) by chromosome segment
510 duplications in Neurospora crassa. *Genetics* **134**, 267–275 (2008).

511 22. M. Ciullo, M.-A. Debily, L. Rozier, M. Autiero, A. Billault, V. Mayau, S. El Marhomy, J. Guardiola, A.
512 Bernheim, P. Coullin, D. Piatier-Tonneau, M. Debatisse, Initiation of the breakage-fusion-bridge mechanism
513 through common fragile site activation in human breast cancer cells: the model of PIP gene duplication from a
514 break at FRA7I. *Hum. Mol. Genet.* **11**, 2887–2894 (2002).

515 23. A. Hellman, E. Zlotorynski, S. W. Scherer, J. Cheung, J. B. Vincent, D. I. Smith, L. Trakhtenbrot, B. Kerem,
516 A role for common fragile site induction in amplification of human oncogenes. *Cancer Cell* **1**, 89–97 (2002).

517 24. M. Marotta, X. Chen, A. Inoshita, R. Stephens, G. Thomas Budd, J. P. Crowe, J. Lyons, A. Kondratova, R.
518 Tubbs, H. Tanaka, A common copy-number breakpoint of ERBB2 amplification in breast cancer colocalizes
519 with a complex block of segmental duplications. *Breast Cancer Research* **14**, R150 (2012).

520 25. M. Marotta, X. Chen, T. Watanabe, P. W. Faber, S. J. Diede, S. Tapscott, R. Tubbs, A. Kondratova, R.
521 Stephens, H. Tanaka, Homology-mediated end-capping as a primary step of sister chromatid fusion in the
522 breakage-fusion-bridge cycles. *Nucleic Acids Res* **41**, 9732–9740 (2013).

523 26. M. Marotta, T. Onodera, J. Johnson, G. T. Budd, T. Watanabe, X. Cui, A. E. Giuliano, A. Niida, H. Tanaka,
524 Palindromic amplification of the ERBB2 oncogene in primary HER2-positive breast tumors. *Sci Rep* **7**, 41921
525 (2017).

526 27. J. J. Bianchi, V. Murigneux, M. Bedora-Faure, C. Lescale, L. Deriano, Breakage-Fusion-Bridge Events
527 Trigger Complex Genome Rearrangements and Amplifications in Developmentally Arrested T Cell
528 Lymphomas. *Cell Reports* **27**, 2847-2858.e4 (2019).

529 28. M. L. Hoang, F. J. Tan, D. C. Lai, S. E. Celniker, R. A. Hoskins, M. J. Dunham, Y. Zheng, D. Koshland,
530 Competitive Repair by Naturally Dispersed Repetitive DNA during Non-Allelic Homologous Recombination.
531 *PLOS Genetics* **6**, e1001228 (2010).

532 29. U. Oggenfuss, D. Croll, Recent transposable element bursts are associated with the proximity to genes in a
533 fungal plant pathogen. *PLoS Pathog* **19**, e1011130 (2023).

534 30. N. T. Umbreit, C.-Z. Zhang, L. D. Lynch, L. J. Blaine, A. M. Cheng, R. Tourdot, L. Sun, H. F. Almubarak, K.
535 Judge, T. J. Mitchell, A. Spektor, D. Pellman, Mechanisms generating cancer genome complexity from a
536 single cell division error. *Science* **368** (2020).

537 31. B. McClintock, The Production of Homozygous Deficient Tissues with Mutant Characteristics by Means of
538 the Aberrant Mitotic Behavior of Ring-Shaped Chromosomes. *Genetics* **23**, 315–376 (1938).

539 32. B. McClintock, The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics* **26**, 234–282 (1941).

540 33. F. Toledo, G. Buttin, M. Debatisse, The origin of chromosome rearrangements at early stages of AMPD2 gene
541 amplification in Chinese hamster cells. *Current Biology* **3**, 255–264 (1993).

542 34. X. Bi, S.-C. D. Wei, Y. S. Rong, Telomere Protection without a Telomerase: The Role of ATM and Mre11 in
543 Drosophila Telomere Maintenance. *Current Biology* **14**, 1348–1353 (2004).

544 35. G. H. Rank, W. Xiao, A. Kolenovsky, G. Arndt, FLP recombinase induction of the breakage-fusion-bridge
545 cycle and gene conversion in Saccharomyces cerevisiae. *Curr Genet* **13**, 273–281 (1988).

546 36. J. A. Hackett, D. M. Feldser, C. W. Greider, Telomere dysfunction increases mutation rate and genomic
547 instability. *Cell* **106**, 275–286 (2001).

548 37. J. Maciejowski, T. de Lange, Telomeres in cancer: tumour suppression and genome instability. *Nat. Rev. Mol.*
549 *Cell Biol.* **18**, 175–186 (2017).

550 38. K. A. Smith, P. A. Gorman, M. B. Stark, R. P. Groves, G. R. Stark, Distinctive chromosomal structures are
551 formed very early in the amplification of CAD genes in Syrian hamster cells. *Cell* **63**, 1219–1227 (1990).

552 39. C. Ma, S. Martin, B. Trask, J. L. Hamlin, Sister chromatid fusion initiates amplification of the dihydrofolate
553 reductase gene in Chinese hamster cells. *Genes Dev.* **7**, 605–620 (1993).

554 40. A. Coquelle, E. Pipiras, F. Toledo, G. Buttin, M. Debatisse, Expression of Fragile Sites Triggers
555 Intrachromosomal Mammalian Gene Amplification and Sets Boundaries to Early Amplicons. *Cell* **89**, 215–
556 225 (1997).

557 41. T. Badet, S. Fouché, F. E. Hartmann, M. Zala, D. Croll, Machine-learning predicts genomic determinants of
558 meiosis-driven structural variation in a eukaryotic pathogen. *Nat Commun* **12**, 3551 (2021).

559 42. M. H. Lendenmann, D. Croll, E. L. Stewart, B. A. McDonald, Quantitative trait locus mapping of
560 melanization in the plant pathogenic fungus *Zymoseptoria tritici*. *G3* **4**, 2519–2533 (2014).

561 43. G. H. Kema, E. C. Verstappen, M. Todorova, C. Waalwijk, Successful crosses and molecular tetrad and
562 progeny analyses demonstrate heterothallism in *Mycosphaerella graminicola*. *Curr. Genet.* **30**, 251–258
563 (1996).

564 44. S. B. Goodwin, T. A. J. van der Lee, J. R. Cavaletto, B. te Lintel Hekkert, C. F. Crane, G. H. J. Kema,
565 Identification and genetic mapping of highly polymorphic microsatellite loci from an EST database of the
566 septoria tritici blotch pathogen Mycosphaerella graminicola. *Fungal Genetics and Biology* **44**, 398–414
567 (2007).

568 45. G. C. Allen, M. A. Flores-Vergara, S. Krasynanski, S. Kumar, W. F. Thompson, A modified protocol for
569 rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nature Protocols* **1**, 2320
570 (2006).

571 46. K. McCluskey, B. W. Russell, D. Mills, Electrophoretic karyotyping without the need for generating
572 protoplasts. *Curr. Genet.* **18**, 385–386 (1990).

573 47. J. Sambrook, D. W. Russell, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory
574 Press, Cold Spring Harbor, N.Y, 3rd ed., 2001).

575 48. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

576 49. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features.
577 *Bioinformatics* **26**, 841–842 (2010).

578 50. C.-S. Chin, D. H. Alexander, P. Marks, A. A. Klammer, J. Drake, C. Heiner, A. Clum, A. Copeland, J.
579 Huddleston, E. E. Eichler, S. W. Turner, J. Korlach, Nonhybrid, finished microbial genome assemblies from
580 long-read SMRT sequencing data. *Nature Methods* **10**, 563–569 (2013).

581 51. Marçais, Guillaume, Delcher, Arthur L., Phillippy, Adam M., Coston, Rachel, Salzberg, Steven L., Zimin,
582 Aleksey, MUMmer4: A fast and versatile genome alignment system. *PLOS Computational Biology* **14**,
583 e1005944.

584 52. P. Rice, Longden, Lan, Bleasby, Alan, EMBOSS: The European Molecular Biology Open Software Suite.
585 *Trends in Genetics* **16**, 276–277 (2000).

586 53. R. Kolpakov, G. Bana, G. Kucherov, mreps: Efficient and flexible detection of tandem repeats in DNA.
587 *Nucleic Acids Res.* **31**, 3672–3678 (2003).

588 54. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The
589 Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

590 55. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in
591 performance and usability. *Molecular Biology and Evolution* **30**, 772–780 (2013).

56. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).

57. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **11**, 1–10 (2016).

58. J. Castresana, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**, 540–552 (2000).

59. L. G. Wang, T. T. Y. Lam, S. Xu, Z. Dai, L. Zhou, T. Feng, P. Guo, C. W. Dunn, B. R. Jones, T. Bradley, H. Zhu, Y. Guan, Y. Jiang, G. Yu, Treeio: An R Package for Phylogenetic Tree Input and Output with Richly Annotated and Associated Data. *Molecular Biology and Evolution* **37**, 599–603 (2020).

60. H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the Tidyverse. *Journal of Open Source Software* **4**, 1686 (2019).

61. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T. Y. Lam, Ggtree: an R Package for Visualization and Annotation of Phylogenetic Trees With Their Covariates and Other Associated Data. *Methods in Ecology and Evolution* **8**, 28–36 (2017).

62. G. Yu, T. T. Y. Lam, H. Zhu, Y. Guan, Two methods for mapping and visualizing associated data on phylogeny using GGTree. *Molecular Biology and Evolution* **35**, 3041–3043 (2018).

63. P. Edge, V. Bansal, Longshot enables accurate variant calling in diploid genomes from single-molecule long read sequencing. *Nat Commun* **10**, 4660 (2019).

Data availability: The genome assembly and annotation for 1E4, 1A5, and 3D7 genome are available at the European Nucleotide Archive (http://www.ebi.ac.uk/ena) under accession numbers PRJEB15648, PRJEB20900 and PRJEB14341. Progeny genomes are available under the project PRJNA645795.

Author contributions: SF and DC conceived the study; SF and UO performed analyses; BAM and DC supervised the work and provided funding; SF, UO and DC wrote the manuscript with input from all co-authors.

Funding information

This work was supported by grants from the Swiss National Science Foundation to BAM (155955), DC (173265) and UO (P5R5PB_225522).

Competing interests

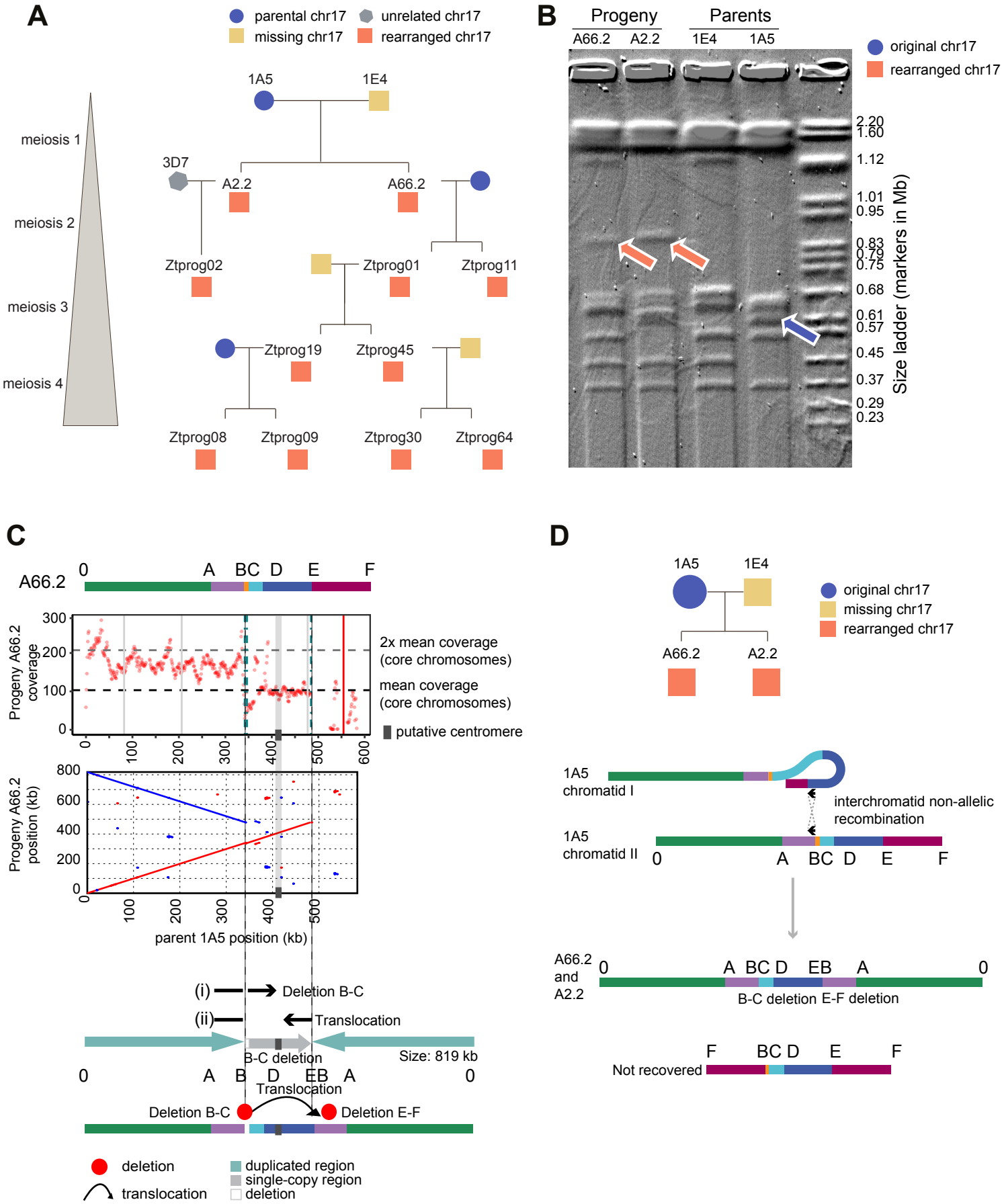
We declare to have no competing interests.

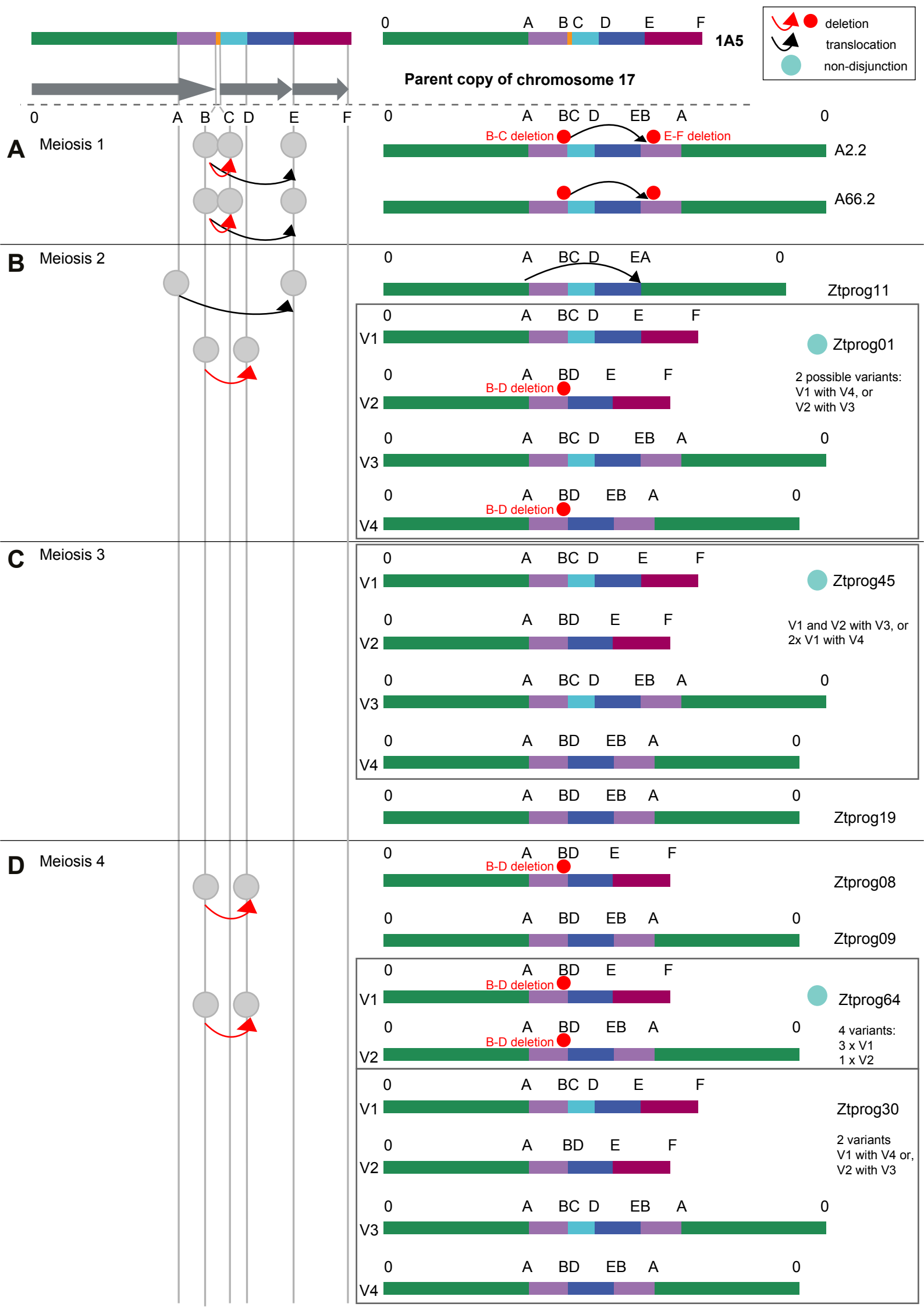
Acknowledgements

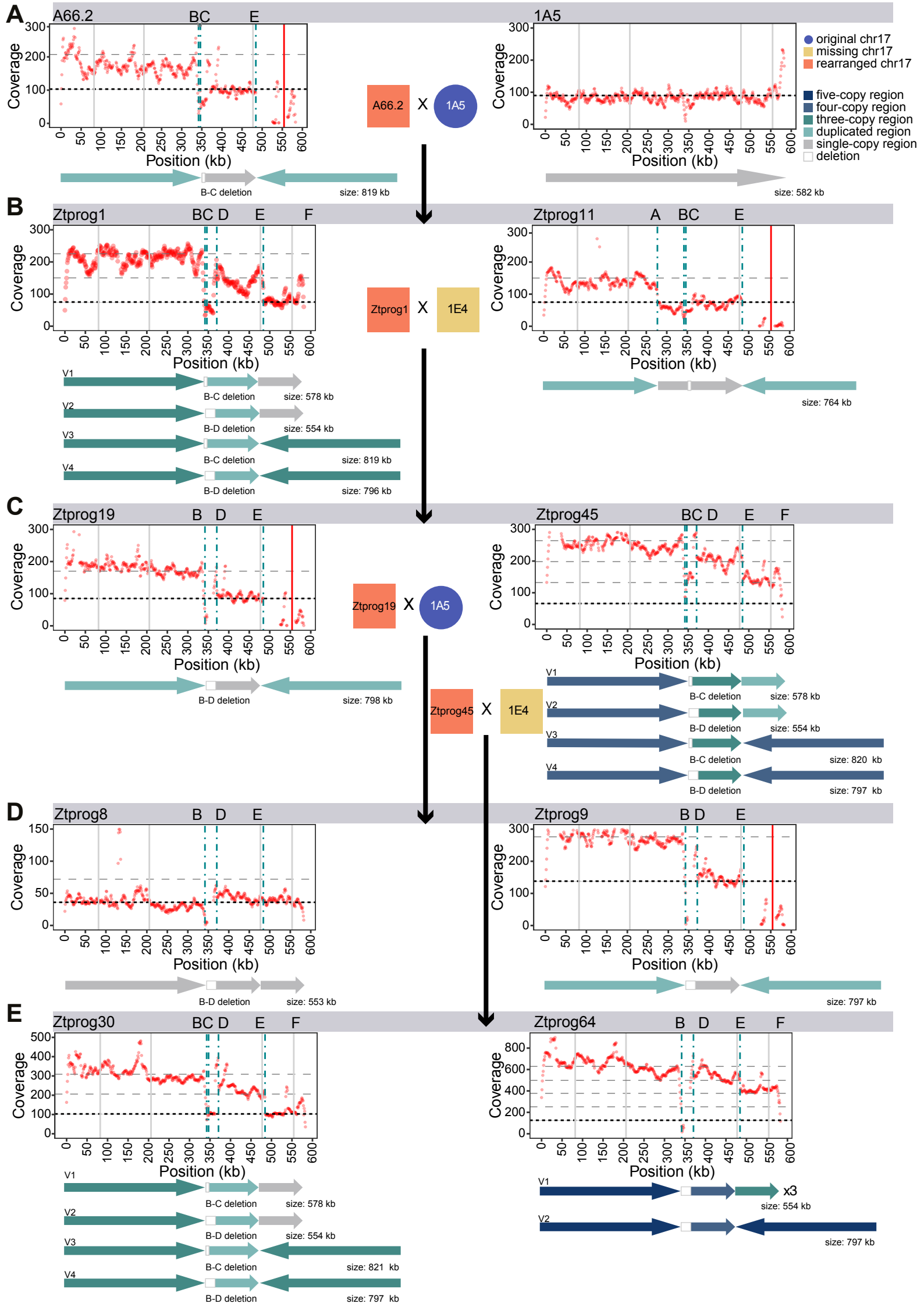
We are grateful for technical support in the laboratory by Marcello Zala. Thomas Badet provided advice on genome sequence analyses.

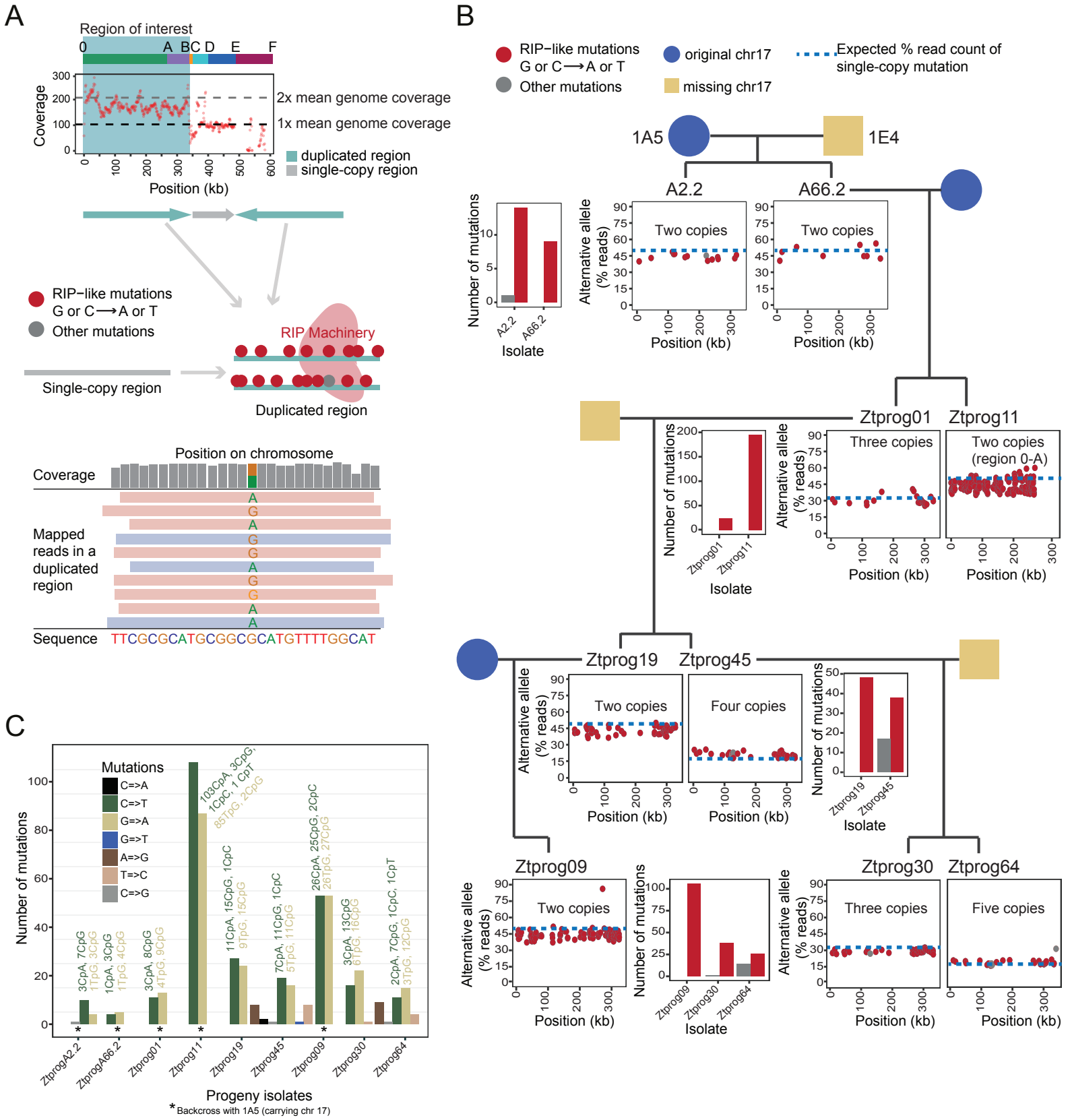
Supplementary Materials

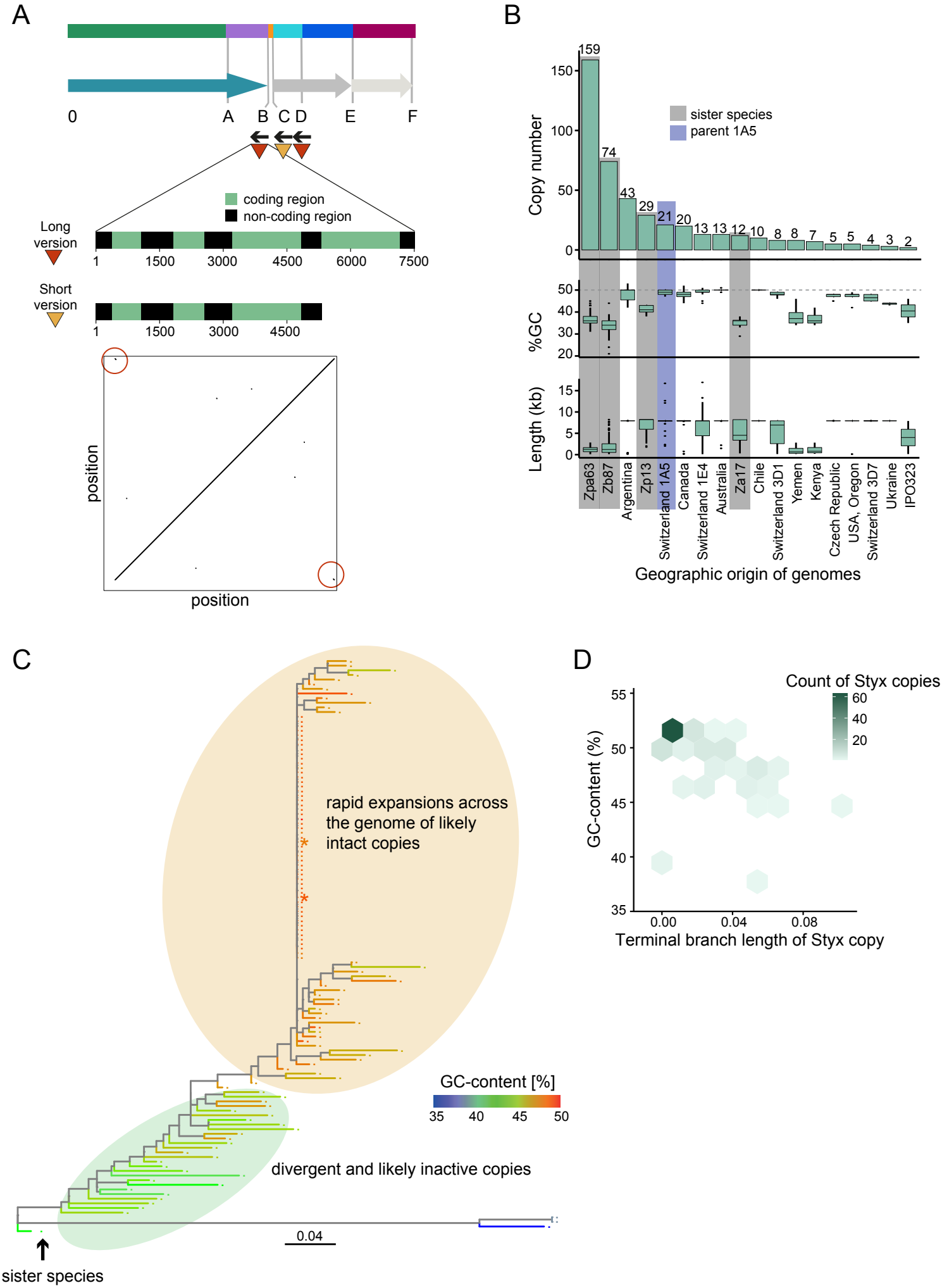
Supplementary Figures S1–S12 and Supplementary Tables S1–S7.


A
parental chr17
unrelated chr17
missing chr17
rearranged chr17
meiosis 1
meiosis 2
meiosis 3
meiosis 4
1A5
1E4
3D7
A2.2
A66.2
Ztprog02
Ztprog01
Ztprog11
Ztprog19
Ztprog45
Ztprog08
Ztprog09
Ztprog30
Ztprog64
B
Progeny
Parents
A66.2
A2.2
1E4
1A5
original chr17
rearranged chr17
2.20
1.60
1.12
1.01
0.95
0.83
0.79
0.75
0.68
0.61
0.57
0.45
0.37
0.29
0.23
Size ladder (markers in Mb)
C
A66.2
0
A
BC
D
E
F
Progeny A66.2 coverage
2x mean coverage (core chromosomes)
mean coverage (core chromosomes)
putative centromere
Progeny A66.2 position (kb)
parent 1A5 position (kb)
(i)
Deletion B-C
(ii)
Translocation
B-C deletion
Size: 819 kb
Translocation
Deletion B-C
Deletion E-F
deletion
translocation
duplicated region
single-copy region
deletion
D
1A5
1E4
A66.2
A2.2
original chr17
missing chr17
rearranged chr17
1A5 chromatid I
interchromatid non-allelic recombination
1A5 chromatid II
A66.2 and A2.2
EB
B-C deletion E-F deletion
Not recovered


deletion
translocation
non-disjunction
1A5
Parent copy of chromosome 17
A
Meiosis 1
B-C deletion
E-F deletion
A2.2
A66.2
B
Meiosis 2
Ztprog11
Ztprog01
2 possible variants:
V1 with V4, or
V2 with V3
B-D deletion
C
Meiosis 3
Ztprog45
V1 and V2 with V3, or
2x V1 with V4
Ztprog19
D
Meiosis 4
Ztprog08
Ztprog09
Ztprog64
4 variants:
3 x V1
1 x V2
Ztprog30
2 variants
V1 with V4 or,
V2 with V3


A
A66.2
BC
E
Coverage
Position (kb)
B-C deletion
size: 819 kb
A66.2
X
1A5
1A5
size: 582 kb
original chr17
missing chr17
rearranged chr17
five-copy region
four-copy region
three-copy region
duplicated region
single-copy region
deletion
B
Ztprog1
BC D
E
F
V1
B-C deletion
size: 578 kb
V2
B-D deletion
size: 554 kb
V3
B-C deletion
size: 819 kb
V4
B-D deletion
size: 796 kb
Ztprog1
X
1E4
Ztprog11
A
BC
E
size: 764 kb
C
Ztprog19
B
D
E
B-D deletion
size: 798 kb
Ztprog19
X
1A5
Ztprog45
BC D
E
F
V1
B-C deletion
size: 578 kb
V2
B-D deletion
size: 554 kb
V3
B-C deletion
size: 820 kb
V4
B-D deletion
size: 797 kb
Ztprog45
X
1E4
D
Ztprog8
B
D
E
B-D deletion
size: 553 kb
Ztprog9
B D
E
size: 797 kb
E
Ztprog30
BC D
E
F
V1
B-C deletion
size: 578 kb
V2
B-D deletion
size: 554 kb
V3
B-C deletion
size: 821 kb
V4
B-D deletion
size: 797 kb
Ztprog64
B
D
E
F
V1
x3
size: 554 kb
V2
size: 797 kb


A
Region of interest
0
A
B
C
D
E
F
Coverage
2x mean genome coverage
1x mean genome coverage
duplicated region
single-copy region
Position (kb)
RIP−like mutations
G or C⟶A or T
Other mutations
RIP Machinery
Single-copy region
Duplicated region
Position on chromosome
Coverage
Mapped reads in a duplicated region
Sequence
TTCGCGCATGCGGCGCATGTTTTGGCAT
B
RIP−like mutations
G or C⟶A or T
Other mutations
original chr17
missing chr17
Expected % read count of single-copy mutation
1A5
1E4
A2.2
A66.2
Number of mutations
Isolate
Alternative allele (% reads)
Two copies
Position (kb)
Ztprog01
Ztprog11
Three copies
Two copies (region 0-A)
Ztprog19
Ztprog45
Four copies
Ztprog09
Ztprog30
Ztprog64
Five copies
C
Mutations
C=>A
C=>T
G=>A
G=>T
A=>G
T=>C
C=>G
Number of mutations
3CpA, 7CpG
1TpG, 3CpG
1CpA, 3CpG
1TpG, 4CpG
3CpA, 8CpG
4TpG, 9CpG
103CpA, 3CpG, 1CpC, 1 CpT
85TpG, 2CpG
11CpA, 15CpG, 1CpC
9TpG, 15CpG
7CpA, 11CpG, 1CpC
5TpG, 11CpG
26CpA, 25CpG, 2CpC
26TpG, 27CpG
3CpA, 13CpG
6TpG, 16CpG
2CpA, 7CpG, 1CpC, 1CpT
3TpG, 12CpG
ZtprogA2.2
ZtprogA66.2
Ztprog01
Ztprog11
Ztprog19
Ztprog45
Ztprog09
Ztprog30
Ztprog64
Progeny isolates
*Backcross with 1A5 (carrying chr 17)


A
0
A
B
C
D
E
F
coding region
non-coding region
Long version
1
1500
3000
4500
6000
7500
Short version
1
1500
3000
4500
position
position
B
Copy number
159
74
43
29
21
20
13
13
12
10
8
8
7
5
5
4
3
2
sister species
parent 1A5
%GC
Length (kb)
Zpa63
Zb87
Argentina
Zp13
Switzerland 1A5
Canada
Switzerland 1E4
Australia
Za17
Chile
Switzerland 3D1
Yemen
Kenya
Czech Republic
USA, Oregon
Switzerland 3D7
Ukraine
IPO323
Geographic origin of genomes
C
rapid expansions across the genome of likely intact copies
GC-content [%]
35
40
45
50
divergent and likely inactive copies
0.04
sister species
D
Count of Styx copies
60
40
20
GC-content (%)
55
50
45
40
35
0.00
0.04
0.08
Terminal branch length of Styx copy