Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal

Consequences of chromosome structure polymorphism in sexual populations

Maria Costa Neves Ferreira da Silva

Dissertação

Mestrado em Biologia Molecular e Genética

2014

Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal

Consequences of chromosome structure polymorphism in sexual populations

Maria Costa Neves Ferreira da Silva

Dissertação orientada pelo Doutor Miguel Godinho Ferreira (IGC) e pelo Professor Doutor Júlio Duarte (FCUL)

Mestrado em Biologia Molecular e Genética

2014

This thesis resulted in the publication of a poster:

Maria Ferreira da Silva, Lília Perfeito, Ana Teresa Avelar, Miguel Godinho Ferreira (2013). ***Testing the chromosome speciation model in fission yeast***. IX Encontro Nacional de Biologia Evolutiva (IX National Evolutionary Biology Meeting), 20th December 2013, Oeiras, Portugal.

Acknowledgements

Agradeço a todos os que contribuíram para a elaboração desta tese.

Ao Miguel, por me ter acolhido no seu laboratório, por me ter confiado o desenvolvimento deste projecto, e por me ter iniciado no mundo da ciência.

Ao Professor Júlio Duarte, por ter aceitado ser meu orientador interno, pela sua disponibilidade, e por me ter ensinado muito do que sei sobre cromossomas.

À Lília, agradeço a disponibilidade, a ajuda, e a paciência. Agradeço ainda as muitas revisões desta tese.

À Teresa, por tudo o que me ensinou sobre leveduras e técnicas. Por me ter ajudado imensamente neste último ano, pelas discussões, pelo companheirismo, e por me ter ouvido. Obrigada ainda por ter lido este manuscrito, e pelas suas sugestões.

To all my colleagues in the Telomeres group, for the help, suggestions, scientific discussions, and, mostly, for welcoming me so warmly in the lab. For the good moments, and the rides late at night, I thank them.

To the people I met at the IGC during this year, for making my stay at the institute so much better. I thank Pol the most.

À minha Tia Isabel, por me ter dado guarida tantas vezes.

Aos meus pais e ao meu irmão. Nem uma palavra desta tese teria sido escrita sem o seu apoio. A eles devo tudo isto.

Finalmente, ao Francisco. Por todo o amor e compreensão.

Abstract

It has been postulated that speciation can occur without geographical isolation. Chromosome rearrangements, such as inversions and translocations, have been proposed as drivers for speciation without isolation (sympatry). In sexual populations, meiosis between structurally different genomes would have high lethality. Recombination between non-linear genomes may produce progeny with imbalanced DNA content. Therefore, only meiotic products which maintain the breakpoints are viable. This causes linkage between the breakpoints, and with alleles of nearby genes. The latest models propose that rearrangements can entrap mutations which accumulate over time near the breakpoints. If recombination leads to epistasis between mutations in linkage with the breakpoints and others which are not, then selection for suppression of recombination can appear. Suppression of recombination will cause further divergence between karyotypes, and ultimately speciation.

An experimental evolution setup was designed to test the chromosome speciation model, using *Schizosaccharomyces pombe*. Cells containing an inversion in one chromosome were mixed with wild type (WT) cells to create sympatric circumstances. Long periods of mitotic adaptive evolution were intercalated with controlled cycles of meiosis, for a total of 500 generations.

We observed that the evolution of sterility was a recurrent phenotype under direct selection caused by sympatric sexual reproduction, appearing in half of the populations. In some populations sterility was fixed. In the remaining, it was maintained as a polymorphism with mating-proficient cells. In this is case, polymorphism for structure and mating type was present, unlike in the former.

In one of the polymorphic populations, we also observed an increase in the frequency of healthy asci (a proxy for meiotic success) produced by heterozygous crosses. Interestingly, this was uncoupled from spore viability, which suffered no changes relatively to the ancestral crosses. Additionally, we did not observe any change in recombination rates.

Our results suggest that the most common solution to meiotic incompatibility in microbial sexual populations in sympatry is sterility. This phenomenon abolishes meiotic fitness depression by completely eliminating sex, leaving karyotypes' fates under exclusive control of adaptive forces. We thus found a new mechanism that may lead to genetic divergence between different karyotypes in facultative sexual populations.

Keywords: chromosome rearrangements, speciation, meiosis, sterility

Resumo

Os genomas eucariotas estão organizados em unidades individuais de ADN compactado, os cromossomas. A conformação e ordem de genes em cada cromossoma são características de cada espécie. Não obstante, é possível encontrar variabilidade na configuração cromossómica de indivíduo para indivíduo dentro de uma espécie. Uma alteração na configuração cromossómica é denominada rearranjo cromossómico. Rearranjos cromossómicos incluem duplicações ou delecções de secções do cromossoma, inversões de segmentos genómicos, ou translocações de segmentos entre cromossomas diferentes. Os rearranjos cromossómicos podem ter consequências drásticas. Em humanos, rearranjos cromossómicos estão normalmente associados a patologias, como cancro. Por exemplo, a leucemia mielóide crónica é causada por uma translocação entre os cromossomas 9 e 22 que resulta numa fusão de genes. Esta fusão leva a uma desregulação da expressão génica normal, e, consequentemente, a uma sobreproliferação celular. No entanto, noutras espécies, como *Drosophila pseudobscura*, os rearranjos cromossómicos estão presentes dentro de populações naturais em estado polimórfico. Sabe-se que, nestas populações, rearranjos cromossómicos estão associados a processos adaptativos, isto é, determinadas estruturas cromossómicas são vantajosas em determinados ambientes. Consequentemente, as frequências de dado rearranjo cromossómico nas populações estão dependentes do ambiente envolvente. Também em leveduras se observou que a estrutura cromossómica influencia a adaptação ao meio, podendo certos rearranjos ser benéficos.

Apesar de poderem ser benéficos na adaptação, os rearranjos cromossómicos apresentam-se sempre como uma desvantagem durante o processo de reprodução sexuada. Durante a meiose I, os cromossomas homólogos emparelham e recombinam. Quando um dos cromossomas homólogos contém um rearranjo, por exemplo uma inversão, o emparelhamento está dependente da formação de uma estrutura em laço. Após a recombinação, dois dos produtos meióticos formados serão letais, pelo facto de conterem duplicações de genes, delecções, e/ou um número errado de centrómeros. Dada a letalidade ou abaixamento de *fitness* resultante do processo, a isto se denomina depressão meiótica. Devido à inviabilidade dos produtos meióticos recombinados, os dois pontos de quebra (*breakpoints*) de um rearranjo estarão sempre em *linkage* entre si, pois não podem ser segregados independentemente sem depressão meiótica.

Dadas estas observações, concluiu-se que os rearranjos cromossómicos podem ser mantidos numa população através de pleiotropia antagonista. Adicionalmente, estas observações levaram a questionar se, a longo prazo, os rearranjos cromossómicos podem ter influência na divergência de populações. O modelo mais aceite de especiação, ou separação de duas populações em duas espécies distintas, diz que a separação ocorre na

presença de isolamento geográfico entre as duas populações. A este fenómeno chama-se especiação alopátrica. No entanto, muitos defendem que a especiação pode ocorrer na ausência de isolamento físico entre populações – especiação simpátrica – e que os rearranjos cromossómicos podem ser catalisadores do processo. O modelo de especiação cromossómica mais recente postula que mutações podem ser acumuladas ao longo do tempo junto dos *breakpoints* de um rearranjo, devido ao *linkage* entre si e consequentemente com os *loci* adjacentes. Estes alelos não serão partilhados com indivíduos que não possuam o rearranjo por causa da depressão meiótica. Este abaixamento no *flow* génico pode levar a uma divergência entre cariótipos. Se, após uma meiose, ocorrer epistasia entre alelos em *linkage* com os *breakpoints* e outros que não estejam, pode criar-se uma pressão selectiva para supressão de recombinação. A supressão de recombinação iria acelerar a divergência e, eventualmente, formar-se-iam barreiras à reprodução, levando à especiação.

Utilizando *Schizosaccharomyces pombe*, levedura de fissão, desenhámos uma experiência para testar o modelo de especiação cromossómica por supressão de recombinação. Durante 500 gerações, duas populações com configurações cromossómicas diferentes (uma com uma inversão e outra sem a inversão) evoluíram em simpatria através de ciclos de meiose (no total 5) intercalados com períodos longos de mitose (100 gerações). Após cada meiose, todas as células vegetativas foram eliminadas com recurso a *snail juice*, uma mistura de enzimas digestivas de caracol, e apenas os esporos resultantes da meiose foram seleccionados para o período seguinte de mitose. Com isto pretendíamos submeter a população total a uma forte selecção para reprodução sexuada, especialmente entre genomas incompatíveis, isto é, rearranjo com não-rearranjo.

Após as 500 gerações de evolução, encontrámos que a solução mais comum para escapar ao cenário de forte incompatibilidade foi o aparecimento de esterilidade. Em pelo menos metade das populações réplica que foram submetidas a este regime de evolução este fenótipo apareceu: em metade destas, a esterilidade fixou-se, sendo que a totalidade dos indivíduos perdeu totalmente a capacidade de se reproduzir sexuadamente, e na outra metade indivíduos estéreis coexistem com os restantes. Ao comparar as populações sexuadas em simpatria com populações de controlo que evoluíram através do mesmo regime mas apenas com indivíduos com a mesma estrutura cromossómica (ou seja, em alopatria), observámos que a esterilidade foi consequência directa da incompatibilidade genómica, pois esta não apareceu em alopatria. Também verificámos que a esterilidade não apareceu noutras populações mistas assexuadas, o que reforça a conclusão de selecção directa. Verificou-se uma perda de variabilidade em alopatria e simpatria, embora esta tenha sido mais acentuada em alopatria; isto porque a pleiotropia antagonista associada à meiose é capaz de manter a variabilidade cromossómica em simpatria durante mais tempo. No

entanto, nas populações onde a esterilidade evoluiu, a estrutura cromossómica fixou-se: o desaparecimento da meiose levou ao desaparecimento da pleiotropia antagonista, e o destino do rearranjo passou a estar dependente apenas do seu *fitness* em mitose. Através de sequenciação do genoma dos clones estéreis, encontrámos mutações pontuais candidatas à causa do fenótipo. A esterilidade, numa das populações estudadas, foi detectada após o segundo ciclo de meiose, o que sugere que a selecção foi bastante forte. De forma a escaparem à meiose, os estéreis ganharam resistência ao *snail juice* utilizado na experiência. A evolução de resistência à selecção por *snail juice* conferiu uma vantagem adicional aos estéreis: ao sobreviver à selecção estas células encontram o nicho vazio e dividem-se mais depressa do que os esporos germinam.

Em *S. pombe*, a meiose dá origem a quatro esporos envolvidos num asco. Esta estrutura denomina-se tétrada. Devido ao facto de o tamanho dos esporos reflectir o seu conteúdo genético (isto é, a quantidade de ADN que herdaram), este tamanho, bem como a sua forma e número, constituem uma *proxy* para o sucesso da meiose. Quando a meiose se dá entre genomas incompatíveis, como quando um dos indivíduos contém um rearranjo, as tétradas produzidas têm morfologias anormais. Numa das populações simpátricas sexuadas analisadas, observámos que, após 500 gerações, os cruzamentos entre indivíduos com rearranjo e sem rearranjo passaram a produzir uma proporção maior de tétradas de 4 esporos, ou tétradas saudáveis. Esta proporção (cerca de 90%) é comparável à proporção encontrada para cruzamentos homozigóticos, e significativamente maior que a proporção em cruzamentos heterozigóticos no início da experiência (cerca de 70%). Isto sugere que durante a experiência teria aparecido um mecanismo que tivesse compensado a depressão meiótica, como, por exemplo, supressão de recombinação. No entanto, encontrámos que a viabilidade meiótica dos esporos na geração 500 se manteve igual à viabilidade no tempo 0. Também não detectámos qualquer evolução nas taxas de recombinação em cruzamentos heterozigóticos. Apesar de tudo isto, o facto de não termos detectado este fenótipo em alopatria ou em populações assexuadas sugere que poderá ter aparecido em consequência da pressão para o cruzamento entre genomas incompatíveis. Permanece, no entanto, por perceber que vantagem este fenótipo confere.

Em conclusão, o modelo de especiação cromossómica permanece por testar, uma vez que as nossas populações escaparam à pressão selectiva ao evoluir a incapacidade de se reproduzir sexuadamente. Isto eliminou a depressão meiótica e a possibilidade de epistasia negativa entre mutações surgidas em diferentes estruturas, sendo portanto benéfica para ambos os cariótipos. Encontrámos pois um novo mecanismo que poderá levar a divergência entre cariótipos em populações microbianas facultativamente sexuadas.

Palavras-chave: rearranjos cromossómicos, especiação, meiose, esterilidade

Index

List of figures

List of tables

Abbreviations

BFB – breakage-fusion-bridge

bp – base pairs

CNV – copy number variation

DMSO – dimethyl sulfoxide

DSB – double strand break

FACS – fluorescence activated cell sorting

FoSTeS – fork stalling and template switching

GFP – green fluorescent protein

Inv – Inverted

LCR – low copy repeat

Mb - megabase

MRN – Mre11, Rad50, Nbs1 protein complex

NAHR – non-allelic homologous recombination

NHEJ – non-homologous end joining

OD_{600} – optical density at 600nm

PCR – polymerase chain reaction

RSA – random spore analysis

SE – standard error

SJ – snail juice

WT – wild type

I. Introduction

It has long been known that biological diversity is not limited to the sequence of the genome, but that there are many other factors which influence individual phenotypes and, by extension, variation within populations and species. A classic example of a source of phenotype variation not encoded in the primary sequence is epigenetics, where chemical modifications occurring on DNA or DNA-associated histones affect gene expression. While epigenetics has been largely explored, other factors contributing to phenotype variation have often been overlooked. One of these is the structure in which the sequences are organized in the genome. Structure itself will in principle lead to variation and diversity by defining the molecular microenvironment surrounding genes and therefore affecting their regulation.

Genomic structure is relevant for all organisms, particularly eukaryotes whose genome is organized into a number of chromosomes. Eukaryotic chromosomes are single identities composed of coiled DNA packed together by histones and other proteins. These DNA-protein complexes define chromatin regions which determine the expression state of genes: euchromatin regions are associated with highly expressed genes, whereas heterochromatin associates with silenced genes. Recently, however, it has become evident that this is not the only way in which chromosomal structure defines regulatory states and gene statuses. Recent high-throughput techniques have made possible the study of the importance of spatial organization of the genome[1,2]. Particular emphasis has been given to intra-chromosomal chromatin interactions as well as chromosome positioning and association within the nucleus. A classic example of how structure influences gene expression is the case of the α-globin genes: to activate expression, the chromosome must acquire a specific folding which allows the spatial proximity of the gene and its distal regulatory elements[3]. These data indicate that chromosomes often reach high-order folding states that allow the transcription of genes or families of genes that share transcription machineries. This also gives clues as to why chromosomes occupy specific territories inside the nucleus.

Genome structure influences gene regulation in several levels. However, the effects of changes in chromosomal configuration remain unexplored, both at the cellular and population levels.

I.1. Causes of chromosome rearrangements

Chromosome rearrangements are changes in chromosomal structure. These changes can be deletions or duplications of genes or genomic segments, translocations, or inversions of chromosomal regions. Chromosome rearrangements usually appear as a consequence of double-stranded DNA breakage at one or more sites – called breakpoints – and subsequent re-joining of the segments producing a different arrangement of the sequence.

Chromosome rearrangements are present within populations and contribute to intra-species genetic variation[4,5,6], and several studies show how chromosome structure influences phenotype and variability[6,7,8]. Rearrangements are, in fact, a predominant feature of cancer[9,10]. It has been suggested that cancer appears as a consequence of the accumulation of small genomic rearrangements and point mutations over time[11,12], or because of a large genomic crisis whereby hundreds of rearrangements occur in a one-off event[13]. The latter is, however, a rare event. It is currently an open question whether, in the majority of cancers, chromosome rearrangements are a cause of malignancy, by interfering with gene regulation or affecting important protein-coding genes, or are a direct consequence of pre-existing genomic instability in malignant or pre-malignant cells. For example, mutations in genes involved in DNA mismatch repair, such as in the case of hereditary nonpolyposis colorectal cancer[14], or homologous recombinational DNA repair (e.g., *BRCA1*)[15] will greatly increase the mutation rate and result in genomic instability, leading to chromosome rearrangements.

Genome instability can also be a consequence of high levels of DNA damage, such as double-strand breaks (DSBs)[16,17]. Improper repair by the cell's repair systems can lead to the generation of chromosome rearrangements (figure 1a)[18,19]. Non-homologous end joining (NHEJ), non-allelic homologous recombination (NAHR), and fork-stalling and template switching (FoSTeS) are some of these processes.

The NHEJ pathway repairs DSBs via protein recruitment. In mammals, the pathway is initiated with the recruitment of the Ku70/80 heterodimer to both ends of the DSB. This complex creates a molecular scaffold promoting the assembly of other NHEJ proteins, such as DNA-PKcs, a protein kinase, and the MRN (Mre11, Rad50, and Nbs1) complex[20], among others. These catalyse the formation of a synaptic complex that brings both DNA ends together, which are then processed in order to be compatible for ligation. This step is thought to be responsible for the loss and alteration of nucleotides during this type of repair and is mediated by both nucleases and polymerases. The ligase IV/XRCC4 then ligates the ends[21].

NAHR is a type of repair that occurs between low-copy repeats (LCRs) originating from previous duplications, and is particularly common in certain genome hotspots[22] such as loci associated with copy number variation (CNV). Repair occurs by homologous recombination between highly similar but non-allelic regions[18] and can amplify CNV regions[23].

FoSTeS is a complex phenomenon that has been described recently[24]. This phenomenon occurs when the replication complex encounters an obstacle in the template DNA, such as DNA modifications, repeats, or secondary structures that stall the fork. Because of this stalling, the 3' end becomes free from its template and invades another fork by aligning with an exposed single strand that shares microhomology. The replication complex switches templates and reinitiates polymerisation in the new fork[23]. This produces a

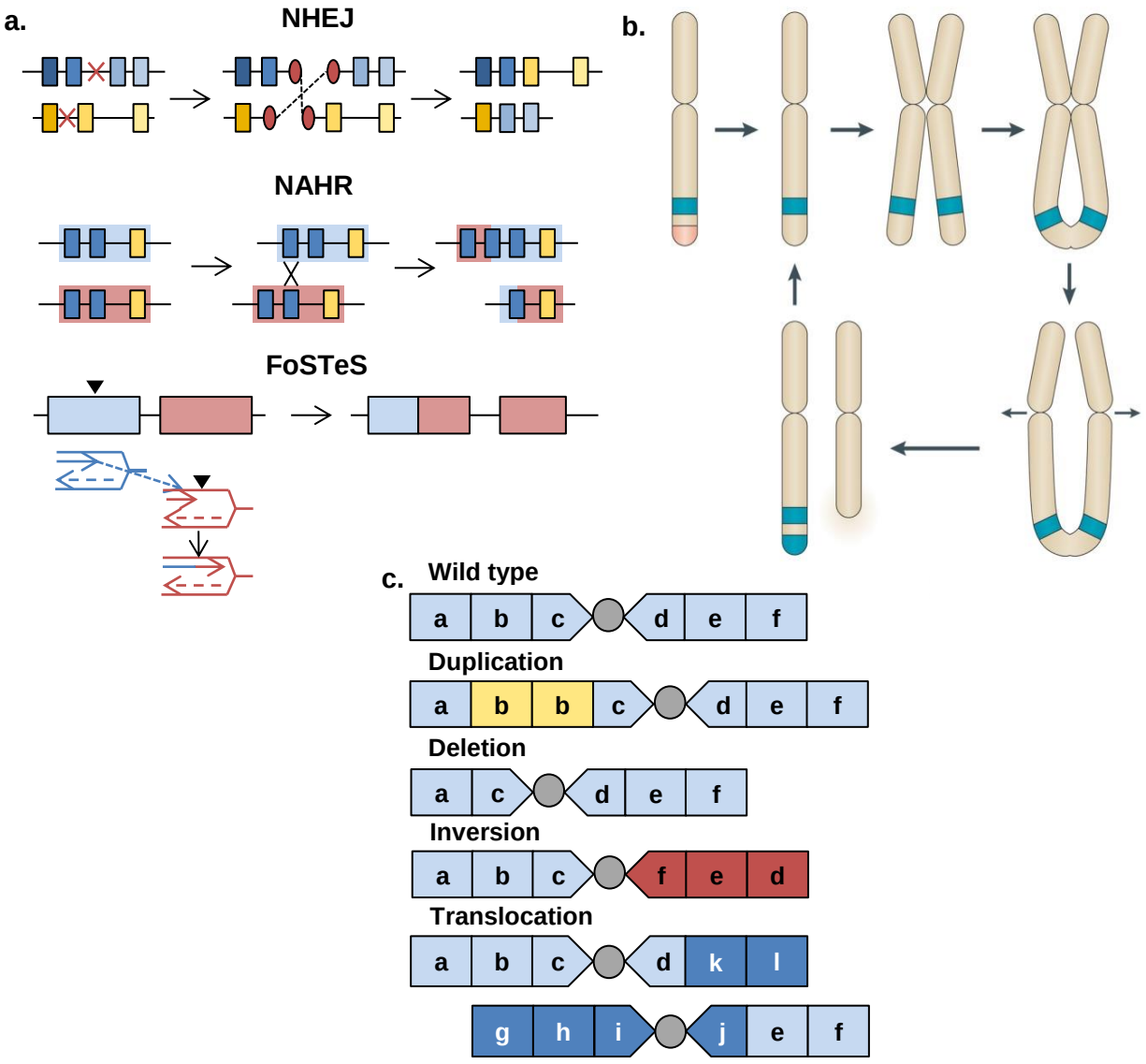


Figure 1. Causative mechanisms and types of chromosome rearrangements. **a.** Hypothesized mechanisms for the formation of chromosome rearrangements involve the action of various DNA repair systems. In NHEJ, DSBs (red crosses) lead to the recruitment of specific proteins (red ovals) which repair lost nucleotides and ligate the strands together; rearrangements occur when strands from different chromosomes are ligated. NAHR results, in this specific case, in a duplication and deletion in two different chromosomes; blue rectangles represent non-allelic repetitive sequences in the genome that recombine. In FoSTeS, the DNA replication complex is stalled due to chemical or structural modifications in the DNA (black triangle), leading to the association of the growing strand (blue dashed line) to a region with high content similarity, which can result in a duplication (as pictured), deletion, inversion, or translocation depending on the region of association and fork position. Based on [25] and [18]. **b.** Simplified scheme of breakage-fusion-bridge (BFB) phenomenon. The telomere (orange) is lost, causing the sister chromatids to fuse and form a bridge after replication. The centromeres are pulled apart in opposite directions (black arrows), and the chromatids break at a random locus, which can cause gene amplification (green). Note the fusion between sister chromatids is a rare event and the BFB cycle can also be initiated by fusions of different chromosomes which have suffered DSBs. Adapted from [26]. **c.** Types of chromosome rearrangements. The pictured inversion is said to be a paracentric inversion since it does not affect the centromere, in contrast with a pericentric inversion; the pictured translocation is a reciprocal translocation as it

involves the balanced exchange of segments between chromosomes. Other types of translocations are Robertsonian translocations (fusion at the centromeres) and unbalanced translocations. Other events of rearrangement include transpositions and aneuploidies. Grey circles represent the centromeres.

signature duplication in one of the daughter molecules, and a deletion in the other. It can also produce inversions or translocations depending on the position of the invaded replication fork. It has been proposed that it may be responsible for several naturally occurring rearrangements, such as the inversion differentiating the chimpanzee chromosome 10 and its homologue in humans, chromosome 12[27]. This rearrangement has the signature duplications at its breakpoints consistent with FoSTeS events (figure 1a).

Genome instability can also be a consequence of telomere deprotection or dysfunction, through a mechanism first described by Barbara McClintock called breakage-fusion-bridge (BFB) cycle[28] (figure 1b): when a telomere breaks off from a chromosome, and if there is no cell cycle arrest, the process of replication will result in two sister chromatids lacking telomeres, which causes them to fuse together at the telomeric site. Fusions can also occur between different chromosomes that have suffered DSBs[29]. A dicentric structure is then formed. During cell division, particularly anaphase, the fused chromatids or chromosomes will form a bridge, with one centromere being pulled towards a direction and the other towards the opposite direction. This causes the chromatids or chromosomes to break apart, but the breakage is random and often occurs at different sites than the fusion site: this results in daughter cells receiving unbalanced chromatids or chromosomes lacking telomeres, which will initiate the cycle again in subsequent cell divisions and cause instability. This kind of genome instability is known to cause chromosome rearrangements, as well as gene amplifications and deletions[30].

Chromosome rearrangements can occur as balanced or unbalanced events (figure 1c). Balanced events include those of inversions and translocations, where the order of genes or segments is altered or shifted, but there is no loss of information. In unbalanced rearrangements, such as deletions or duplications, the amount of information contained in the DNA is altered – lost and doubled, respectively – and these are generally largely deleterious. In balanced events, the immediate effects will depend on where the breakpoints are located: if the breaks occur in coding regions, for example, their expression will be disrupted which can have dramatic consequences.

I.2. Genetic and evolutionary effects of rearrangements

One of the most striking examples of chromosome rearrangements affecting phenotype in humans is the case of the Philadelphia chromosome[31], wherein a reciprocal translocation between chromosomes 9 and 22 results in a fusion gene between the *BCR* and *ABL1* genes. This creates a state of deregulation of the cell cycle, because *ABL1*, a proto-oncogene,

becomes constitutively expressed thus influencing cell proliferation and differentiation. This set of events ultimately results in cancer, particularly chronic myelogenous leukaemia[32]. Besides the example of the Philadelphia chromosome in humans, there are other cases of chromosome rearrangements causing cancer: for example, murine plasmacytomas result from dysregulation of the *c-myc* gene which is caused by a translocation between chromosomes 12 and 15[33]. Another example is Ewing's sarcoma in which 85% of cases are caused by a translocation between chromosomes 11 and 22, fusing genes *EWS* and *FL1*. *FL1*, when fused to *EWS*, activates the Ras pathway and leads to overproliferation of erythroblasts[34].

Chromosome rearrangements do not always have deleterious consequences. Several natural populations, particularly fly populations, are polymorphic for chromosomal structure. In 1938, Dobzhansky described populations of *Drosophila pseudoobscura* which differed in their chromosomal structure. These populations were maintained in their respective geographical locations, but overlapping between them, creating polymorphic zones[5]. Decades later, it was observed that the changes in frequency of these different chromosomal rearrangements along with the populations seemed to be connected with environmental fluctuations. It was then concluded that these variations were clinal, and that the different chromosome rearrangements had influence on the overall fitness of the flies and hence on the process of adaptation. This would explain the maintenance of their polymorphic state[7]. In *D. mediopunctata,* it was found that several inversions in chromosome 2 vary seasonally, with certain polymorphisms being favoured in the warm seasons and others during the cold ones. The In(3R)P inversion of *D. melanogaster* also shows seasonal and latitudinal frequency fluctuations[35]. In plants, one example is the yellow monkeyflower, *Mimulus guttatus,* where certain chromosome configurations are associated with perennial or annual phenotypes[36]. In the budding yeasts *Saccharomyces cerevisiae* and *S. uvarum*, the recurrent evolution of a specific rearrangement in independent lines between both species was reported recently. This rearrangement created an interspecific fusion junction at both *MEP2* gene copies, encoding an ammonium permease, that conferred higher fitness than the parentals in ammonium limitation[37].

Clues as to why chromosome rearrangements are so common in natural populations, while they seem to be mostly deleterious in humans, come from Avelar et al., 2013[6]. This study, done in the fission yeast *Schizosaccharomyces pombe,* showed that rearrangements have effects even if not directly involving coding regions. They demonstrated that: I. the expression of genes located near the sites of the breakpoints is affected by their genomic contextualization; II. the presence of a rearrangement changes the transcription levels of genes across the entire genome, even of genes located in chromosomes not involved in the rearrangement. The phenotypic differences, based on these results, seem to be a consequence

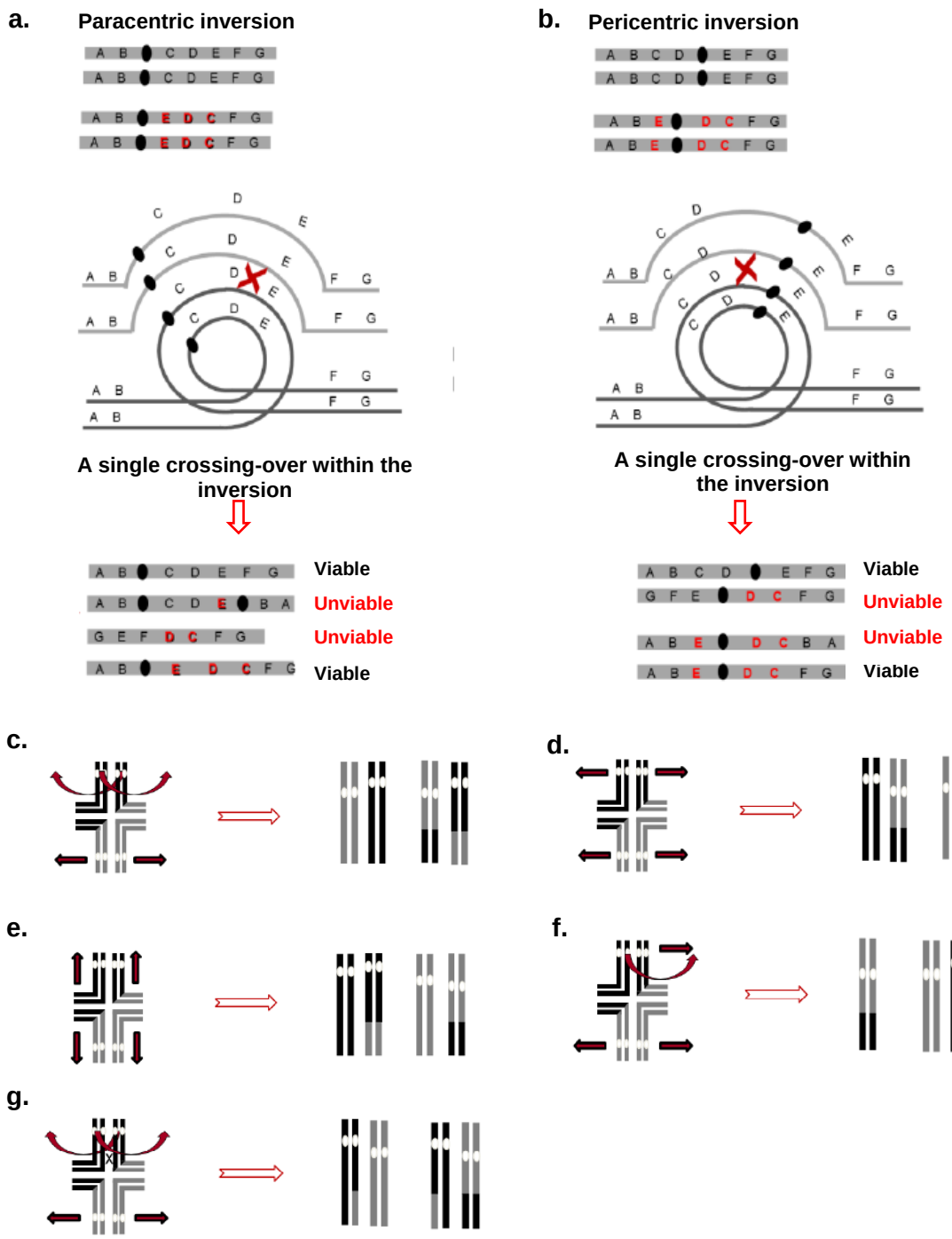


Figure 2. Meiotic pairings with rearranged chromosomes. a. To undergo recombination, an inverted chromosome and its homologue form a particular structure. In the case of a paracentric inversion (not including the centromere), a single crossing-over inside the inverted region will produce two viable and two unviable meiotic products, one of which is dicentric while the other is acentric. **b.** In the case of a pericentric inversion, two unviable products are also produced, containing duplications and deletions. Black circles represent centromeres. Adapted from [38]. **c.** A tetravalent structure is formed in the presence of a balanced translocation. Segregation can be of different types, in this case alternate segregation where all products are viable. **d.** Adjacent type-I segregation, all products are inviable. **e.** Adjacent type-II segregation, all products are inviable. **f.** 3:1 disjunction, products are inviable. **g.** alternate segregation with chiasmata, inviable products. White circles represent centromeres. Adapted from [39].

of repositioning and reordering of entire chromosome domains. These results highlight the importance of gene positioning and strengthen the conclusion that genome architecture within the nucleus is important and chromosomes are not randomly organized in it. This study also showed that chromosome rearrangements, even when not directly affecting coding regions, are a selectable trait. Firstly, they observed that there is wide structural diversity in natural isolates of fission yeast, despite their nucleotide diversity being very low. The amount of natural variability for chromosome structure in *S. pombe* was also quantified by Brown et al., 2011[40] who concluded that 20% of natural isolates show some type of rearrangement. This suggests that either chromosome structure in this species is neutral or associated with beneficial phenotypes. Avelar et al., 2013[6] found that the latter is very likely, by measuring fitness values in many different genome structures. They found that rearrangements can confer either an advantage or a disadvantage in mitotic propagation in different environments. These results point to the importance of chromosome rearrangements in adaptation, possibly because of the changes in gene expression they cause.

Despite the fact that rearrangements can be beneficial in adaptive evolution, they are largely deleterious during the process of meiosis. During meiosis I, homologous chromosomes align in the equatorial plane of the cell and form bivalent structures through the formation of chiasmata. At this point, homologous recombination occurs, with the exchange of alleles between chromosomes. When a chromosome contains a rearrangement, homology is not complete, creating a problem in alignment. To allow full synapsis and subsequent recombination, special structures are formed: loop-like structures (inversions) and tetravalent structures (translocations)[39]. The resolution of these structures produces cells that will have deletions, duplications, or a wrong number of centromeres (in the case of inversions) and are therefore deleterious and/or lethal (figure 2).

The biggest consequence of the phenomenon described in figure 2 is the death of most offspring resulting from sexual reproduction between two structurally different genomes. This could lead to selective pressure to eliminate sexual reproduction, compensate for the meiotic defects, or eliminate the rearrangement from the population. However, because of the advantages certain rearrangements can provide during asexual reproduction, they can be maintained in polymorphic state in the population by antagonist pleiotropy[6]. Additionally, selection could also lead to increased divergence between the karyotypes, because the meiotic depression caused by heterozygous pairings is counter-selected[6,41,42].

I.3. Chromosome rearrangements as an evolutionary driver: the chromosome speciation hypothesis

The question of whether speciation can occur without geographical separation has been asked for a long time, and chromosome rearrangements have been proposed as a

driving mechanism of such phenomenon[43]. There are several models which attempt to explain how different chromosome configurations in a population could lead to speciation in sympatry. The earliest model, the hybrid-underdominance model, proposes that rearrangements act as a catalyst of the speciation process due to their meiotic defects in heterozygocity[42,44,45]. During meiosis, when the chromosome containing a rearrangement encounters its non-rearranged homologue, recombination will result in defective products containing deletions and duplications of genes (figure 2). Taking into account this mechanistic difficulty, the heterokaryotypic hybrids will exhibit reduced fitness in relation to the homozygous types. It is said the hybrids are therefore underdominant, because in this case selection will be disruptive and favour the homozygous genotypes. Selection will also favour mutations that reduce the probability of intercrossing, which could eventually lead to reproductive isolation and species separation.

White's stasipatric model[42] postulates that a chromosome rearrangement can initially be established in a small peripheral population by genetic drift. Low fitness of hybrids will result in limited gene flow between this population and the ancestral one, allowing the evolution of reproductive barriers by mutation and selection and thus speciation *in situ*. White attributed the multiplication of species of Australian grasshoppers of the subfamily Morabinae to this phenomenon[42]. However, along with other hybrid-underdominance models, White's stasipatric model faces a paradoxical difficulty. These models assume strong hybrid underdominance, but if such underdominance existed, how would the chromosome rearrangement be fixed in the peripheral population in the first place? It could be argued that hybrids would have an only slight decrease in fitness, allowing for expansion of the rearrangement into the ancestral population. However, this raises another problem. In this situation, selection would not be strong enough to fix mutations that reduce intercrossing and lead to reproductive isolation[46].

In the last decades, another class of models, the suppressed-recombination models, has arisen[41,43,47]. Briefly, suppressed-recombination models postulate that the effects of chromosome rearrangements on recombination, and not on hybrid fitness, are responsible for the evolution of reproduction barriers. As previously described, recombination between a rearranged and a non-rearranged chromosome results in defective products and thus in unviable offspring. Due to the unviability of recombined chromosomes, the breakpoints of the rearrangement remain in linkage in viable offspring. Because breakpoints are not segregated independently, recombination cannot occur freely in their vicinities[43,44]. Selection would favour the fixation of mutations that reduce recombination because of the disadvantage in recombination between rearranged chromosomes. Moreover, suppressed-recombination models suggest that mutations and new alleles can appear in linkage to the breakpoints and not be diluted in the population because recombination in the presence of rearranged

chromosomes is reduced[43]. This accelerated differentiation, caused by the chromosome rearrangement, allows the accumulation of incompatibilities near the breakpoints between the two genetic backgrounds. A balanced number of crossing-overs between rearranged and non-rearranged chromosomes can still occur and allow gene flow, despite being greatly reduced. If alleles away from the breakpoints show negative or positive epistasis with alleles linked to the breakpoints, than the region of low recombination can be further extended.

One of the first suppressed-recombination models was suggested by Coluzzi[41,48] to explain speciation events within the mosquito species complex *Anopheles gambie*. Coluzzi proposed that marginal or peripheral populations, exposed to different environmental conditions in relation to the main population, adapt divergently, and that these adaptive alleles would be protected from recombination by chromosome rearrangements. Coluzzi's model, and similar ones that followed[49], is a twist on the Dobzhansky-Muller model[50]. The Dobzhansky-Muller model explains how speciation can occur if two populations become geographically isolated and evolve incompatible alleles (Bateson-Dobzhansky-Muller incompatibilities), while resolving at the same time the paradox of hybrid underdominance and selection. It assumes, however, that geographic isolation is required and that speciation will occur only in allopatric circumstances. Suppressed-recombination models propose that chromosome rearrangements can act as a physical barrier against gene flow, much like geographical separation between populations.

The Navarro-Barton model[51,52] (figure 3) is the most recent model and proposes a theoretical framework to explain the mechanism by which chromosome rearrangements could contribute to speciation. They consider the situation of a population divided into two demes with a certain degree of gene flow between them. If a favourable mutation appears in either one of the demes and migrates to the other, then it will rapidly be fixed in both unless an incompatible allele (through epistasis) is already common in the receiving deme. If this situation results in decreased gene flow, an equilibrium with different incompatible alleles fixed in the different demes may arise and, with it, a genetic barrier to reproduction. Because it can be assumed that the fixation of these new alleles is due to divergent adaptation in both demes, the model conservatively considers uniformly advantageous alleles only. The newly arisen genetic barrier, according to the model, can result in post-zygotic isolation because gene flow in loci in linkage with the incompatible alleles diminishes, creating a snowball-like effect that can lead to the appearance of new barrier alleles. In summary, the barrier will progressively strengthen until full reproductive isolation evolves. Any mechanism that decreases gene flow between subpopulations may be considered as plausible to favour the described speciation process. A chromosome rearrangement may therefore initiate this phenomenon because it traps large genomic fragments against recombination. The model

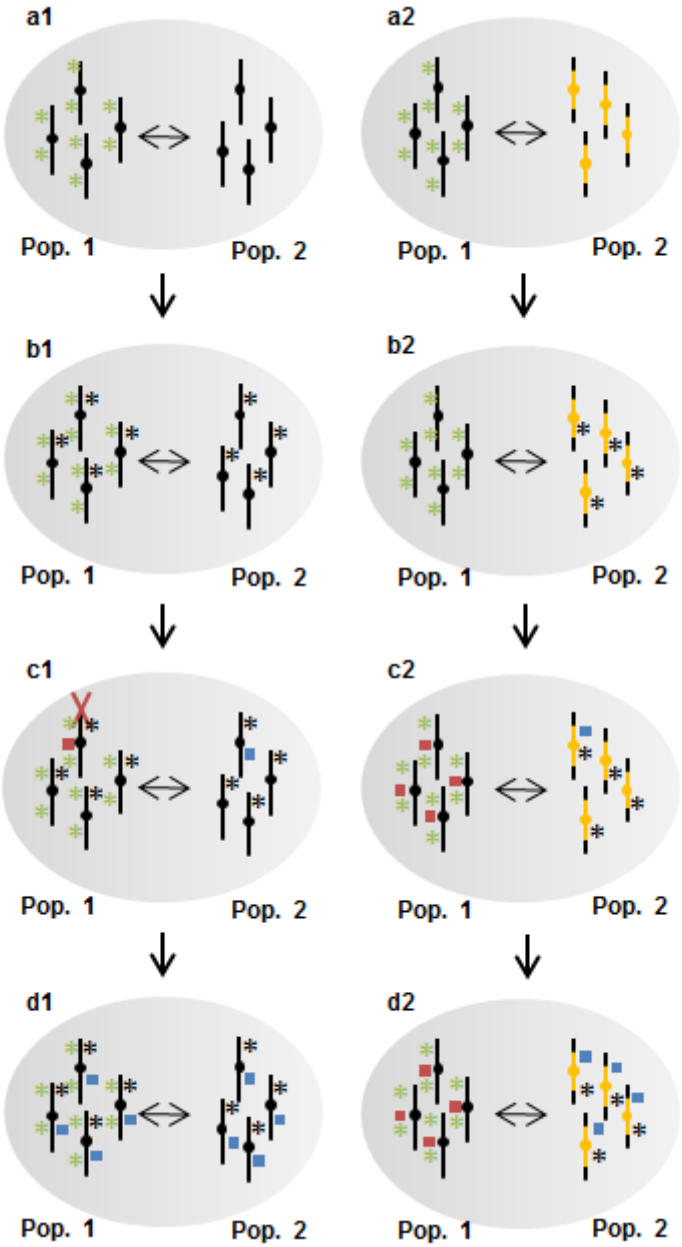


Figure 3. The Navarro-Barton model of chromosome speciation. The starting point of this model is a fixed chromosomal configuration difference between two parapatric populations. A comparison between two scenarios is presented: in one situation, there is no chromosome rearrangement; in the other, a pericentric inversion (in yellow) is fixed in one of the demes. A beneficial allele occurring in collinear regions in one of the populations (black asterisk) will spread and fix inside the deme (b1) or be contained in it if it is trapped by, in this case, an inversion (b2). As such, its flow to the other deme will be delayed in relation to alleles outside the inversion because these recombine freely (not graphically represented). If the delay is long enough, new mutations in the other deme may arise (c2, red square) which may be incompatible with the allele represented by the black asterisk (negative epistasis). These new beneficial mutations will fix in population 1 and will not be able to flow into population 2 because in the hybrid zone both alleles cannot coexist (c2). In contrast, in the absence of a chromosome rearrangement (c1), the red allele will be immediately eliminated from population 1 when it appears because the black allele had spread into it beforehand. In d2, because the chromosome rearrangement is present, new beneficial alleles linked to it (blue square) can appear and contribute to the strength of the barrier

because they are not shareable due to suppressed recombination in the rearrangement, contrary to what happens in scenario d1. Adapted from [44].

successfully combines suppression of recombination and epistasis to explain how chromosome speciation could occur (figure 3).

Theoretical data from the Navarro-Barton model shows that, in the presence of a chromosome rearrangement, the more loci are involved in the barrier the stronger it gets and the more the effective migration rate between demes decreases, even when there's counter-selection against the rearrangement. In comparison with a simulation depicting a similar isolation process between demes without any chromosome rearrangements, they have shown that the presence of a rearrangement actually increases the strength of the barrier[51].

Speciation with suppressed recombination has been observed by Rieseberg et al., 1995 in the wild sunflowers *Helianthus petiolaris* and *H. annuus*[43,53]. By tracking loci in naturally occurring hybrids of the two species, the authors found that in rearranged sections of the chromosomes, very large chromosome blocks behave as a unit in linkage because of lower recombination rates[54]. The biggest example of suppressed recombination and rearrangements leading to chromosome divergence is probably that of the sexual chromosomes in mammals[55,56]. The X and Y chromosomes pair up during meiosis but only in a small conserved homologous region. A series of inversions has led to the establishment of suppression of recombination between the chromosomes and allowed their divergence.

Despite the indication and indirect evidence that chromosome rearrangements can lead to divergence between populations, there is no formal proof or empirical observation that rearrangements are actually capable of triggering an event of speciation without geographical isolation.

II. Aims

The work presented in this thesis is an extension of the project whose results are presented in Avelar et al., 2013[6]. The authors of the project had previously found that chromosome rearrangements are a common feature in *S. pombe* natural strains, and that they are not necessarily neutral in adaptive evolution, and can be beneficial or deleterious, depending on the environment. Moreover, they also found that, despite the meiotic fitness depression in sexual heterozygous crossings, chromosome rearrangements can be maintained in populations through antagonistic pleiotropy. The authors hypothesized, based on these results, that chromosome speciation could be thus established in a medium-term evolution experiment in structurally polymorphic sexual populations, according to the suppressed recombination hypothesis. Using the model organism *S. pombe*, the experiments described in this thesis had the following aims:

i. to investigate the medium-term consequences of the presence of genome rearrangements in populations;
ii. to test the chromosome speciation hypothesis;
iii. to investigate whether natural selection can compensate for the meiotic defects of heterozygous pairings, when chromosome rearrangements are present.

The work presented in this dissertation is an analysis on evolution experiments conducted over the year of 2010 by Ana Teresa Avelar, PhD. Ana Teresa Avelar also performed the work in the allopatry experiment populations and sequencing data processing. Lília Perfeito, PhD contributed for the elaboration of table 1.

III. Methods

Strains, media, and drugs. Strains used in this work are presented in table S1. All media were prepared as described[57,58] and treated with ampicillin at a concentration of 100µg/ml. Cells were grown in YES (Yeast Extract with Supplements) rich liquid medium at 32ºC with shaking and aeration unless otherwise stated. Minimal medium utilized was PMG (Pombe Glutamate) with supplements added as necessary for auxotrophy testing, except adenine auxotrophy, where YE (Yeast Extract) medium was used. For sexual reproduction, ME (Malt Extract) medium was used, supplemented with lysine hydrochloride (225 mg/l). YFM (Yeast Freezing Medium) was utilized to freeze cells at -80ºC. For phleomycin selection, phleomycin (Invivogen) was added to molten YES solid medium to a concentration of 100µg/ml, and plates freshly prepared.

Clone sampling. For genotype analysis, populations were grown at several timepoints from 96-well microplates of the original evolution experiment (stored at -80ºC) to YES solid medium with a microplate replicator and grown at 32ºC. Cells from each population were then grown overnight and subsequently plated in YES solid medium at the appropriate dilutions. Isolated colonies (n=48) from each population and timepoint were picked and grown overnight in 96 deep-well plates. Plates were centrifuged for 3min at 800*g* and the cell pellet was suspended in YFM in a new 96-well microplate appropriate for freezing at -80ºC. The isolated clones were then plated to YES solid medium with a microplate replicator and grown at 32ºC for genotyping.

Genotyping. Because chromosome structure is in linkage with a fluorescent tag in our work, structure genotype was visually identified on a stereomicroscope Stereo Lumar.V12 (Zeiss). Mating type and breakpoint conformation were assessed by colony PCR. Cells grown previously in solid YES medium were mixed with Z buffer (2,5mg/ml of zymolyase, 1,2M sorbitol, 0,1M sodium phosphate pH 7,4, and 1:10 of lysing enzymes 100mg/ml) and incubated at 30ºC for 30min followed by 5min at 95ºC. 2µl of the lysate were added to the PCR mix with 10x DreamTaq Buffer with 20mM $MgCl_2$ (Thermo Scientific), 10mM DNTP mix, 10µM of appropriate primers (table S2), DreamTaq 5U/µl DNA polymerase (Thermo Scientific), and dH_2O milli-Q to final volume. An initial cycle of 5min at 94ºC was performed on the reaction mix, followed by 30 cycles of 30s at 94ºC, 30s at 55ºC, and 3min at 72ºC, and a final extension of 10min at 72ºC. PCR products were resolved in 1% (m/v) agarose in 1x TAE buffer and stained with RedSafe™ Nucleic Acid Staining Solution (iNtRon Biotechnology). Both mating type and breakpoint conformation are assessed by band size: h+ produces a 987bp band, and h- a 729bp band; *lys3* and *his1* breakpoints in the inversion

are respectively 3000bp and 1500bp, while in the control they are 2500bp and 2000bp. Auxotrophies were determined by replica-plating cells onto appropriate media with a microplate replicator.

Induction of mating. Cells were grown overnight, and subsequently grown again to exponential phase on the next day. They were then pelleted at 800*g* for 3min, and resuspended in PMG liquid medium. The wash was repeated two more times. Cells were then resuspended in 100µl of PMG. To achieve sexual reproduction, 5µl of cells of a mating type were mixed with 5µl of a strain with the other mating type in an ME plate. Plates were incubated at 25ºC for 2-3 days. Presence of tetrads was confirmed visually by microscopy, and for each cross 100-150 tetrads were counted and categorized (according to [59]). Mating efficiency was qualitatively assessed visually as well.

Tetrad dissection. Mating was induced as described. From each cross, cells and tetrads were taken and resuspended in 50µl YES liquid medium, and then inoculated in a YES solid plate as a drop at the periphery of the plate. A line across the plate was then formed by letting the drop slide, and allowed to dry. Tetrads were manually picked out of the line with a micromanipulator equipped with a glass needle (Singer), and incubated at 32ºC for 6 hours to allow ascus breakage. Tetrads were then dissected by separating the spores, and the plate was again incubated at 32ºC. After 3-5 days, number of colonies was scored (each spore will or will not give rise to an isolated colony). Meiotic viability is given by the number of colonies observed over the number of colonies expected.

Random spore analysis. Mating was induced as described previously. Cells and tetrads were picked from ME plates and dissolved in 1ml of dH_2O milli-Q. 20µl of a 10% solution of *Helix pomatia* extract (Pall Life Sciences) were then added, and the mixes were incubated at 25ºC with shaking overnight. Mixes were then washed three times in dH_2O milli-Q. The number of spores was assessed using a haemocytometer, and 200-300 spores were plated on YES plates. Meiotic viability is given by the number of colonies over the number of plated spores.

Snail juice resistance assessment. Cells were grown for 2-3 days in ME solid plates at 25ºC, picked, and suspended in dH_2O milli-Q. OD_{600} was measured in a Multiskan GO Microplate Spectrophotometer (Thermo Scientific) and the number of cells was estimated according to a calibration line. *Helix pomatia* extract (Pall Life Sciences) was added to the cells at the concentration of 1:100. The mixes were incubated overnight at 32ºC with shaking, washed three times with dH_2O milli-Q, and then plated at the appropriate dilutions and

volumes on YES solid medium. After incubation at 32ºC, the surviving colonies were counted.

Yeast cells transformation. Cells were grown overnight, and then diluted 1:100 and grown until exponential phase. They were then pelleted and washed in 1ml of LiAc-TE (0,1M lithium acetate, 10mM Tris pH 7,5, 1mM EDTA). After centrifugation, cells were suspended in 100µl of LiAc-TE. 10µl of 10mg/ml sonicated salmon sperm DNA (Stratagene), previously treated for 5min at 95ºC and cooled in ice for another 5min, and 10µl (about 1µg) of DNA for transformation were added. The mix was incubated for 5min at RT. 260µl of LiAc-TE-PEG (LiAc-TE plus 40% PEG4000) were then added. To allow transformation, the mix was incubated an additional hour at 32ºC. Afterwards, a volume of 43µl of dimethyl sulfoxide (DMSO) was added and mixed gently, followed by a 7min heat shock at 42ºC. Cells were pelleted for 1min at 800*g*, and washed once with YES liquid medium. To allow recovery, they were then incubated overnight. The following day, cells were plated on appropriate selective media. Positive colonies were streaked onto fresh selective media plates, confirmed for DNA insertion by PCR, and grown for subsequent storage at -80ºC.

Construction of a recombination marker. An intergenic region in chromosome I starting at nucleotide 2600293 (halfway from each breakpoint) was chosen. Primers were designed to amplify this region (primers A-forward and D-reverse, both 20nt long, table S2). Two additional primers that anneal in the middle of the region were also designed, primers B-reverse and C-forward. These are composed by 20nt of homology to the genome (3') and 20nt of homology to a bleMX6 resistance cassette (5'). Genomic DNA template was obtained by using the smash-and-grab DNA extraction protocol as described previously[60] with modifications. Two PCR reactions were set up as follows: 5x Phusion® HF Buffer (New England Biolabs), 10mM DNTP mix to a final 0,2mM concentration, 10µM of primers A-forward and B-reverse (or primers C-forward and D-reverse – table S2), Phusion® High-Fidelity 2000U/ml DNA polymerase (New England Biolabs), 3µl of genomic DNA template, and dH_2O milli-Q to final volume. Mixes underwent an initial cycle of 4min at 98ºC, followed by 30 cycles of 30s at 98ºC, 30s at 50ºC, and 1min at 72ºC, and a final extension of 5min at 72ºC. The reactions produced two products, product AB and product CD. Presence of these products (AB 349bp and CD 307bp) was confirmed by gel electrophoresis. Products were then purified with the Wizard SV Gel and PCR Clean-up System (Promega), according to manufacturer instructions. To obtain DNA template of the gene to insert, a bleMX6 cassette was amplified using standard primers F2 and R1[61] (table S2) from plasmid pFA6a-bleMX6. The plasmid was a kind gift from Antony Carr (U. Sussex). The PCR program used was described in an earlier section. Another PCR mix was set up as follows: 5x Phusion® HF

Buffer, 10mM DNTP mix to a final 0,2mM concentration, 50ng of products AB and CD, 10µM of primers A-forward and D-reverse, Phusion® High-Fidelity 2000U/ml DNA polymerase, 50ng of template cassette, and dH_2O milli-Q to final volume. The program utilized consisted of 4min at 98ºC, 5 cycles of 15s at 98ºC, 30s at 40ºC, and 3min at 72ºC, plus 25 cycles of 15s at 98ºC, 30s at 55ºC, and 3min at 72ºC, and a final extension of 5min at 72ºC. The obtained product, a fragment of 1661bp consisting of the bleMX6 cassette flanked by two homology regions, was confirmed and purified as previously, and used to transform cells as described. The gene is inserted in the genome by homologous recombination with the corresponding homology region. Insertion was confirmed by PCR, with primers annealing outside the targeted region (inserted: 2045bp, not inserted: 945bp).

Recombination rates measurement. Mating between bleMX6^S and bleMX6^R strains was induced as described. Random spore analysis protocol was then applied to the crosses, and after spore germination spores were checked for fluorescent tag, and tested for phleomycin resistance by replica plating. Recombination rate is given by the number of recombinant spores divided by the number of total spores that germinated for each cross.

Whole-genome sequencing and analysis. Genomic DNA was extracted using the ZR Fungal/Bacterial DNA MiniPrep kit (Zymo Research) and purified with Millipore Membrane Filters of mixed cellulose esters, hydrophilic, pore size 0,025µm (Merck Millipore), and sequenced in a MiSeq desktop sequencer (Illumina) at an average coverage of 40X. Data was filtered using the following pipeline: BamQC (paired-end trimming), BreSeq (mutation analysis), and FastQC. (quality control). Mutations were manually analysed to discard false positives by using Tablet[62], and confirmed by PCR and cycle sequencing, as described in [38].

IV. Results

IV.1. Selection for sex promotes maintenance of chromosomal diversity but is not sufficient to conserve mating type polymorphism

We set up an experiment to ask whether chromosome rearrangements can lead to an event of speciation, or if natural selection would be able to compensate for the meiotic defects caused by chromosome rearrangements instead. To assess it, the experiment required two populations with different chromosome structures evolving together and reproducing sexually. In this work, we used *S. pombe* as a model, and from the 10 genetically engineered strains created in Avelar et al., 2013[6], we chose a strain containing a long inversion in chromosome I and the respective control, which carries the wild type conformation of the same chromosome. The inversion involves the portion of the chromosome between the *lys3* and *his1* genes (about 3,8Mb), which constitute the breakpoints, also present in the wild type (WT). WT cells were tagged with mCherry fluorescent protein, while inverted cells expressed green fluorescent protein (GFP). The fluorescent tag gene was inserted close to the *his1* breakpoints in both genomes (figure 4a). The experiment consisted in rounds of meiosis intercalated with long periods of asexual propagation, as depicted in figure 4b. Specifically, four different subpopulations of cells were mixed at equal frequencies of 0,25 each – WT h+ cells, WT h-, inverted (Inv) h+, and Inv h-, where h+ and h are the two *S. pombe* mating types – and allowed to propagate by mitosis for 100 generations. After this period of time, one cycle of meiosis was induced and the resulting spores separated from the population by killing vegetative cells (figure 4b). The procedure was repeated during 500 mitotic generations (and 5 rounds of meiosis). The relative frequency of each strain was measured by fluorescence activated cell sorting (FACS) every 15 generations.

In parallel with the main experiment, two other experiments were set: one in which the four genotypes evolved exclusively by mitosis (asexual sympatric populations), and another one where wild type cells and inverted cells evolved sexually but separately (allopatric populations). With these we aimed to distinguish which effects of chromosome structure polymorphism were due to sex and which were due to coexistence of incompatible genomes.

The medium used in this experiment has the particularity of containing two distinct carbon sources with different predicted metabolic pathways: maltose and raffinose. We expect two fitness optima in this medium and the potential for divergent adaptation during the 100 generations of adaptive evolution.

Figure 5 shows the structure genotype frequencies over time for 13 replicate populations in sexual (5a) and asexual (5b) circumstances. Unexpectedly, we observed a common phenomenon in all replicate sexual populations: after the first round of meiosis, the

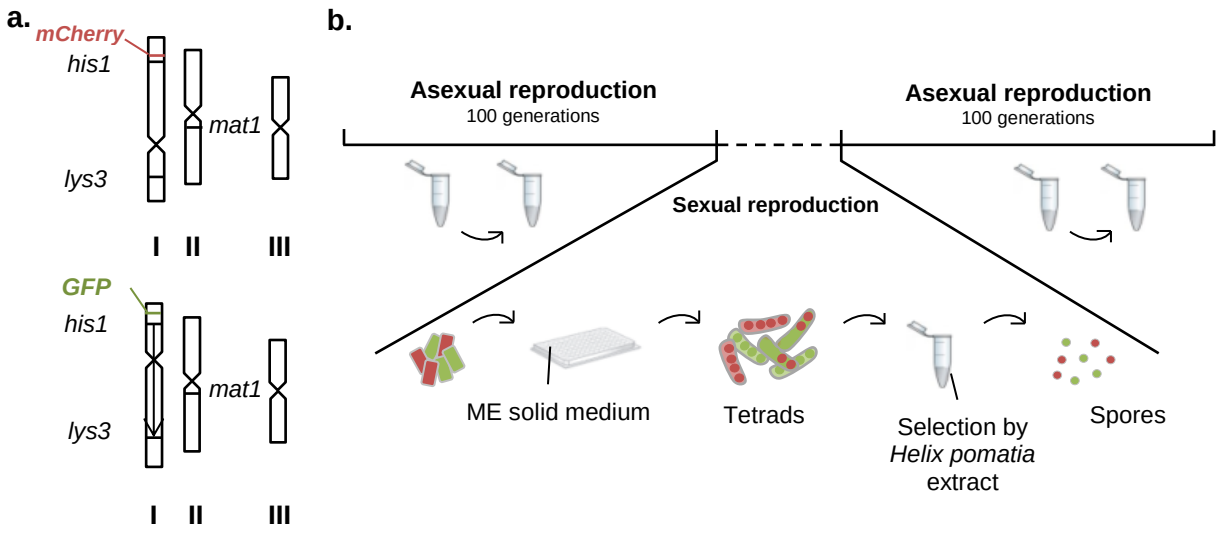


Figure 4. Strains and experimental set-up. **a.** Genomes of the strains used in the main experiment. Top, genome of the WT strains, with mCherry inserted near to the *his1* breakpoint. Bottom, genome of the Inv strains, containing the inversion in chromosome I between genes *lys3* and *his1*, and GFP inserted near to the *his1* breakpoint. The *mat1* gene determining mating type is in chromosome II in both genomes. **b.** Experimental setup for the evolution of sympatric sexual populations. For allopatric populations, the setup is similar, but the genotypes evolve separately and only mate with their equals. For sympatric asexual populations, no intermediate step of sexual reproduction was present.

Inv genotype greatly increased in frequency, after consistently decreasing for about 50 generations. Concomitantly, in asexual conditions we also observed this initial decrease in frequency of the Inv genotype in mitosis after a stable period. In these populations, the Inv went extinct after 300 generations. This could be attributed to an initial advantage of the WT genotype relatively to the Inv one. This was likely the case, as we have seen the recurrent increase in mitosis of the WT h+ genotype and the simultaneous decrease of the WT h-, in sympatric populations, allopatric populations, and asexual sympatric populations (figure 5c and data not shown). These observations suggested that the average fitness of WT was roughly the same as that of Inv, but that WT h+ was much more represented while WT h- was close to extinction at this point. This was confirmed by clone sampling of representative populations (exemplified in figure 5c, population B8 corresponding to the red trajectory in 5a). We chose population B8 hereafter as our main focus of study for two main reasons: I. it maintained polymorphism (see ahead), and II. its steady trajectory (figure 5a, in red) suggests the evolution of compensatory mechanisms. In the sexual experiment, the introduction of meiosis led to the rescuing of the Inv genotype because of recombination. We hypothesized that the mating type frequencies imbalance, due to differences in fitness, was responsible for the 100-generation frequency increase of the Inv genotype after meiosis. From the data resulting from clone sampling of selected populations, we constructed a simple

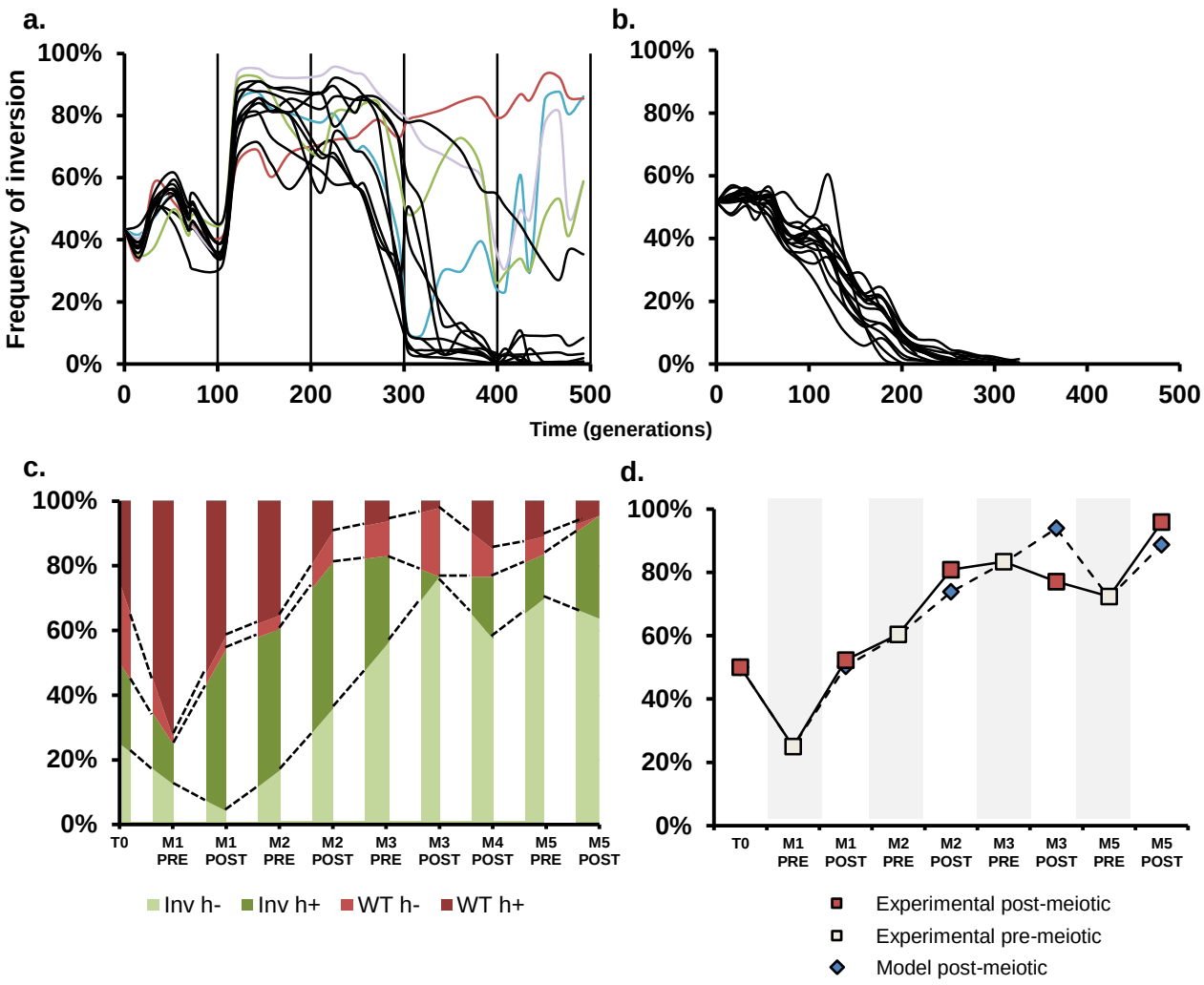


Figure 5. Population trajectories throughout time. a. Trajectories of 13 replicate populations evolved through intercalated cycles of mitosis and meiosis, taken by FACS measurements. Vertical lines represent the points at which meiosis happened. At 100 generations, a steep increase in the frequency of the inverted genotype happened, which was corroborated by our model. In colour, the four populations from which clones were isolated: red – B8, blue - B6, purple – B11, and green – C1. **b.** Trajectories of 15 replicate populations evolved exclusively by mitosis, taken by FACS measurements. By 300 generations, the inverted population went extinct in all the replicates due to the advantage of the WT population **c.** Genotype frequencies throughout time for population B8 (red in 5a) measured by clone sampling in all points pre- and post-meiosis, except the 4th pre-meiotic point. Each discrete coloured bar represents the proportion and frequency of each genotype for the given timepoint. At T0, each genotype is at 0,25 of frequency. **d**. Comparison of experimental data taken from clone sampling throughout time of population B8 and predicted trajectory. Because there are no data from the 4th pre-meiotic point, the 4th post-meiotic frequencies could not be calculated. The model successfully predicts the frequency change of chromosome structure genotype post-meiosis. Grey bars represent points for which the data is exclusively experimental.

model where mating is random between Inv and WT, the advantageous allele(s) is unlinked from structure, and heterozygous crosses yield a lethality rate, in order to assess the consequences of the initial changes in frequency (see supplementary notes for

explanation of model construction). We fitted this model to the data and were able to successfully explain the observed increase in the frequency of the Inv genotype immediately post-meiosis, as exemplified in figure 5d with a representative population (B8, 5c and red trajectory in 5a). This was a consequence of the fact that the WT h+ strain mostly crossed with cells containing the inversion ($P_{WTxWT} = 0{,}01$ vs $P_{WTxI} = 0{,}09$, where P is the probability of mating given by multiplying genotype frequencies, see supplementary notes) due to their relative frequencies at the time of the cross (figure 5c, "M1 PRE"). A proportion of crosses occurred between Inv cells ($P_{IxI} = 0{,}01$), and homozygous crosses yield a very small level of lethality. This brought the Inv subpopulation above 50% of frequency. We later confirmed that most crosses occurred between incompatible genomes in the first meiosis by tracing an allele only the WT h+ possessed at the beginning of the experiment, the *leu1+* allele: all isolated clones resulting from this round of meiosis were descendants from this strain (data not shown). Note that these data refer to population B8, one of the representative populations we studied (figure 5c). The differences in the steepness of the frequency increase between the populations will mostly depend on the exact frequencies of each genotype in the 100 generations prior to meiosis.

Figure 5a shows that, while the inversion was lost in some sexual populations, meiosis was crucial in maintaining the structural diversity in some other, particularly the ones

Table 1. Genotype frequencies of 13 sexual sympatric populations after 500 generations. The population trajectories in colour in figure 5a are highlighted in this table. Whole-population PCR was employed as a preliminary test to check for polymorphism. In the cases where both mating types were observed – the highlighted populations – clone sampling and genotyping was performed to obtain the values below. Further testing of polymorphism was done by plating a whole population sample in ME medium. Populations were considered non-polymorphic if no tetrads were present.

	Wild type *h+*	Wild type *h-*	Inverted *h+*	Inverted *h-*	Tetrads in ME
B6	**27%**	**2%**	**0%**	**71%**	**YES**
B7	100%	0%	0%	0%	NO
B8	**4%**	**0%**	**33%**	**63%**	**YES**
B9	88%	0%	12%	0%	YES
B11	**27%**	**11%**	**7%**	**56%**	**YES**
B12	0%	96%	0%	4%	YES
C1	**0%**	**83%**	**9%**	**9%**	**YES**
C2	99%	0%	1%	0%	NO
C3	99%	0%	1%	0%	NO
C5	100%	0%	0%	0%	NO
C6	0%	73%	0%	27%	NO
C7	100%	0%	0%	0%	YES
C8	91%	0%	9%	0%	YES

whose trajectory is shown in colour. Large frequency fluctuations have occurred much later, particularly but not exclusively during the meiotic step, which suggests the appearance of new mutations. Moreover, different replicate populations had different fates. We sought to assess the frequencies of the four genotypes by the end of the experiment. For this we ran a whole-population PCR on all replicate populations to first confirm the presence of both mating types. Strikingly, 9 out of 13 populations seemed to have lost one mating type during the experiment when tested by PCR: despite the critical role of sexual reproduction in maintaining structural diversity, it was not sufficient to preserve the maintenance of both mating types in many populations. 3 of these 9 populations still produced tetrads when plated in mating medium, and so were considered polymorphic. In total, after two rounds of testing, 6 out of 13 populations lost one mating type. However, four populations still maintained detectable polymorphism. From these, clones were isolated and genotyped. Table 1 shows the breakdown of genotype frequencies for all the replicate populations.

Overall, selection for sexual reproduction led to the loss of one of the mating types in many populations, which suggests that despite strong selection for meiotic spores, meiosis was counter-selected.

IV.2. Sterility is selected in sympatry

The observation of the loss of mating types led us to question whether the presence of an inversion may favour asexual reproduction, and so we investigated whether it had happened in allopatry. Figure 6 shows a comparison between sympatric and allopatric populations. Because of the likely advantage of the WT h+ strain, which we also observed in the allopatry experiment, and no alternate mating counterpart, we expected that the WT h- strain would be lost in most, if not all, replicate populations. This was in fact the case (data not shown). As such, we only included the Inv allopatric populations in the mating type polymorphism comparison in figure 6.

In both scenarios, allopatric and sympatric, close to half of the populations lost mating type polymorphism and consequently meiosis ceased to happen in these populations at one point in the experiment. Chromosome structure seems to have no direct influence over mating type loss (figure 6b and 6d).

After finding that meiosis was counter-selected, we tested whether the sympatric populations that maintained mating type polymorphism were still mating. Surprisingly, we found that 3 out of the 13 populations had completely lost the ability to undergo sexual reproduction (figure 6c shows the total of populations where sterility was found, where these 3 are included). Even more strikingly, all allopatric populations had retained their capacity to mate, despite the early loss of one mating type in the allopatric experiment as well (figure 6f). These results show that coexistence between meiotically incompatible genomes not only led

to mating type loss and counter-selection of sex, but also to direct selection of sterility. We also tested for the asexual populations' ability to mate and found that they had retained it. This strengthens the conclusion that selection for sex between incompatible genomes was directly responsible for the evolution of sterility.

Moreover, we also found that in 3 of the polymorphic populations, B6, B8, and B11 (table 1), sterile cells coexisted with mating-proficient ones. The total number of populations where sterility has evolved may be higher than our estimate, because we cannot exclude the existence of sterile clones in non-polymorphic mating-proficient whole populations (table 1). In total, 6 out of 13 sympatric populations were found to have evolved sterility (figure 6c).

By going to intermediate time points in the experiment and testing cells for their mating capacity throughout time, we found that sterility appeared in population B8 before the second meiosis, in the WT h- background. However, this new subpopulation of sterile cells went rapidly extinct, probably because meiosis was induced immediately after that point. We found sterility again after the second round of meiosis, in the Inv h+ background (the same genetic background where we found it at the end of the experiment). The mutation for sterility was possibly already in the Inv background before meiosis – just below the limit of detection by sampling – and somehow survived the spore selection by resistance or chance, unlike the WT h- sterile cells. Clone sampling and intra-crossing from populations B6, B8, and B11 (blue, red, and purple in figure 5b and table 1, respectively) revealed that, even when mating type loss had not occurred in the entire population, sterile cells coexist with non-sterile cells (figure 5c, B8). In population B8, sterile individuals represent roughly 15% of the population, and half of the total Inv h+ subpopulation.

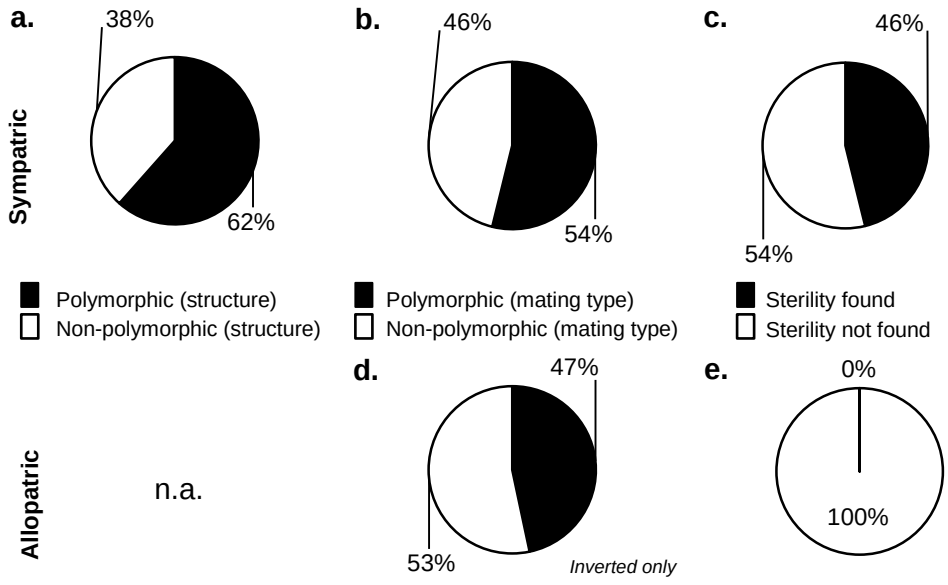


Figure 6. Characterization of sympatric (top) and allopatric (bottom) populations at generation 500. Proportion of polymorphic populations in relation to non-polymorphic concerning **a.** chromosome structure**, b.** and **d.** mating type. **c.** and **e.** proportion of populations where sterility was found to not found. n.a. = not applicable.

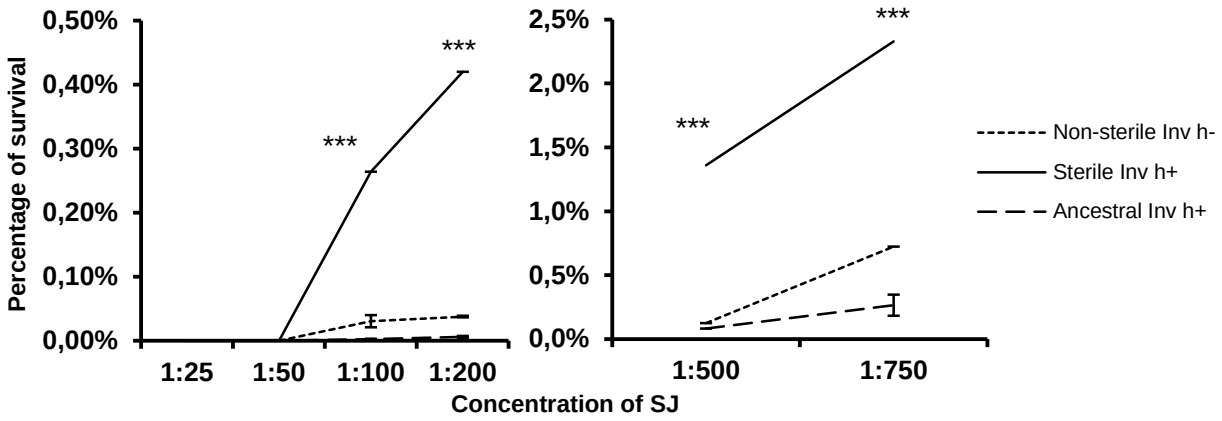


Figure 7. Snail juice resistance. Percentage of survival of two genotypic and phenotypically different clones from population B8 compared to the ancestral strain after being subjected to 6 concentrations of SJ enzymes. In the original experiment, the concentration used to kill vegetative cells was 1:100. Initial number of cells mixed in the SJ solutions for all strains was 6 to $8x10^4$ cells/µl. Error bars represent 2SE of three replicates. A two-tailed T-test was used to assess statistical significance for sterile-non-sterile and sterile-ancestral comparisons, with $p<0,001$ for all (***).

Because meiotic spores were selected in the experiment after inducing mating by adding to the cells a mix of digestion enzymes (snail juice, SJ), the only way sexual reproduction could have been counter-selected was by parallel selection of resistance to these enzymes. Thus we sought to assess the correlation of sterility with SJ resistance. We found that in sympatry sterility directly correlates with resistance to snail juice. We tested snail juice resistance in phenotypically different clones from population B8 and compared them to the ancestral and found that the sterile clone exhibited about 10 times more resistance than the non-sterile and about 20 times more than the ancestral clones with the same chromosome structure at the concentration used in the experiment (figure 7 and figure S1).

IV.3. Selection for sex leads to an increase in healthy asci resulting from meiosis between incompatible genomes

From the previous observations, we concluded that meiosis was counter-selected due to the cost of crossing different chromosome structures. If that was the case, then clones still able to mate were either on the way to being extinct, or compensated for the meiotic defects. We tested that by crossing the evolved clones from population B8 that were still able to mate and estimating their meiotic success.

In *S. pombe*, meiosis occurs after the fusion of two haploid cells of different mating type. After meiosis II, four spores develop inside an ascus, with the spore walls forming around the nuclei. The size of the spore is determined by the quantity of DNA it receives, which means that a successful meiosis will result in four equally sized healthy spores[59]. As such, we utilized spore size and shape as a proxy for meiosis health, according to [59]. A remarkable observation we made when crossing Inv h- and WT h+ clones from generation

500 was that this combination produced much more healthy 4-spore tetrads than the heterozygous ancestral cross (approximately 90% versus approximately 70%, see figure 8a), being closely similar to the homozygous ancestral cross (90% frequency of 4-spore tetrads). Figure 8 shows the plotted data in the form of distributions of 4-spore tetrad frequencies. We observed an increase in the number of 4-spore tetrads in certain heterozygous pairings only, which suggested different evolved genotypes in the same structural background coexisting in the population. Pooling the 4-spore tetrad values of all sympatric heterozygous crosses resulted in a shifted distribution relatively to the ancestral and allopatric heterozygous crosses.

We found that this phenotype evolved in sympatry in the mitotic interval between the second and third meiosis (figure 8b). At this point ("M3 PRE" in figure 5c) all genotypes were present in population B8. However, only heterozygous crosses of WT h+ with Inv h- type presented the phenotype, which means the trait first appeared in one (or both) of these backgrounds; crosses of evolved cells with ancestrals suggest it may have appeared in both as a co-evolutionary trait as these values are similar to ancestral crosses values (figure S2) and that it is not due to new rearrangements. After the third round of meiosis, we did not detect evolved heterozygous crosses presenting the ancestral values anymore (figure 8b).

We sought to understand if this had been a consequence of selection for mating between two different chromosome structures and so we compared these results to the heterozygous evolved cross in allopatry. We performed heterozygous whole-population crossings between allopatric populations and were not able to find a higher proportion of 4-spore tetrads. Allopatric heterozygous evolved crosses yielded a proportion of 4-spore tetrads which was not distinguishable from the heterozygous ancestral cross (figure 8a).

To test if the increase in 4-spore tetrads was accompanied by an increase in spore viability, we performed meiotic viability tests using different methods: tetrad dissection and random spore analysis (RSA). We performed these tests using extreme clones of the distribution: the one with highest fraction of 4-spore tetrads and the one with lowest. Tetrad dissection gives us the meiotic viability associated with these 4-spore tetrads, while RSA gives us the global meiotic viability, for all types of produced tetrads. The increase in 4-spore tetrads could be due to an increase in the number of spores with the correct DNA content. Hence, we expected this would result in an increase in spore viability. However, we found that meiotic viability given by tetrad dissection (figure S3) or RSA (figure 8d) remains the same when compared to the heterozygous ancestral crosses, just below 50%. These results suggest that spore viability is somehow uncoupled from DNA segregation in meiotic division.

Decreased meiotic viability would be an indicator of the evolution of a post-zygotic isolation barrier, as predicted by the chromosome speciation hypothesis. No evidence for decreased meiotic viability, in relation to the ancestral crosses, was found in sympatry or allopatry for the amount of generations tested.

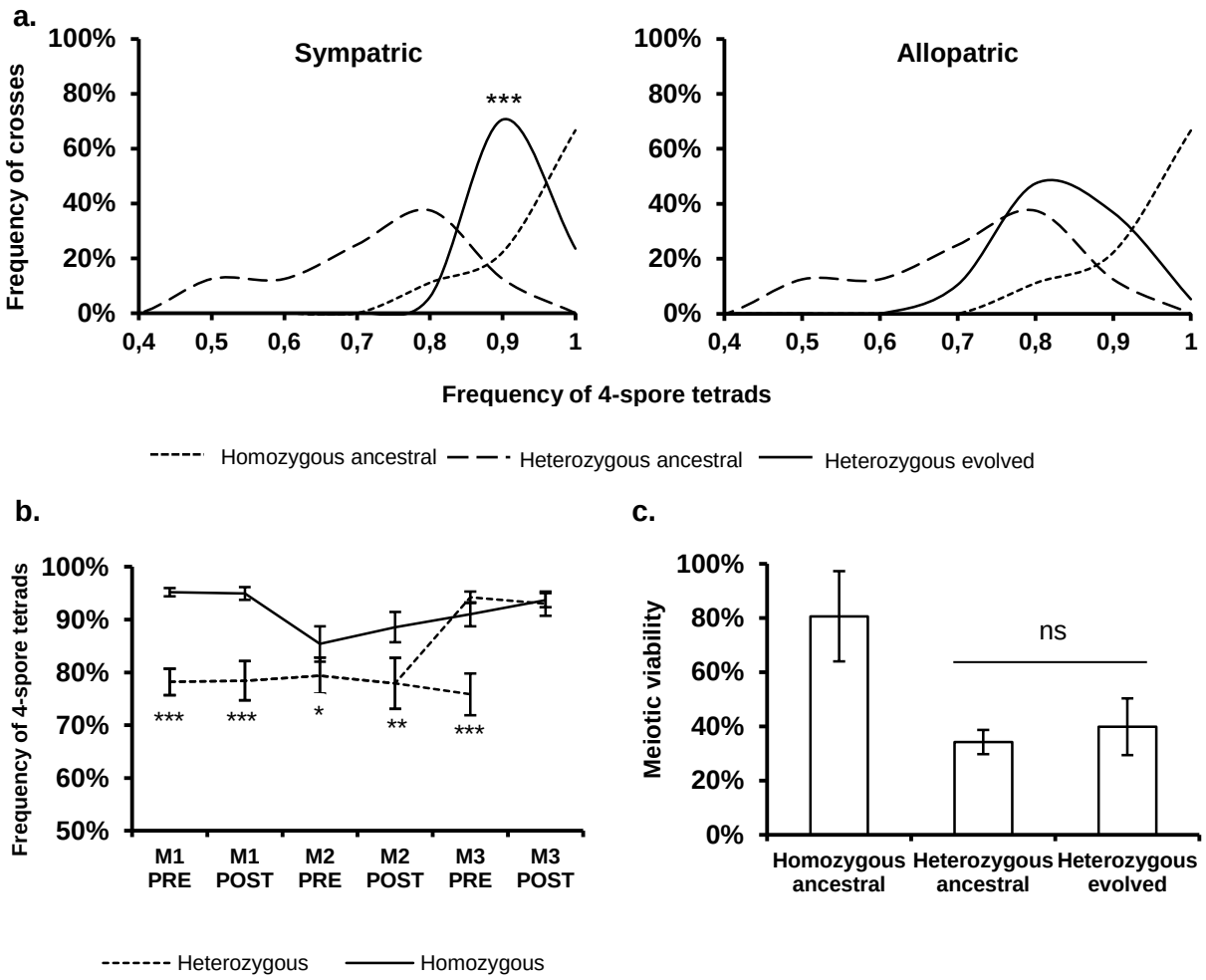


Figure 8. Characterization of meiosis in sympatry after 500 generations. a. Distributions of the frequency of 4-spore healthy tetrads in homozygous ancestral crosses, heterozygous ancestral crosses, and heterozygous evolved crosses in sympatry (population B8) and allopatry. X-axes represent the proportion of 4-spore tetrads in bins of 0,1, while the Y-axes represent frequency of crosses. The distributions plot the data referring to 20-50 crosses, where the ancestral crosses were replicated several times and the evolved distributions are a sum of crosses between different clones in both cases, and thus represent the average of the population. Statistical significance between the heterozygous ancestral and heterozygous evolved crosses was assessed with the Kolmogorov-Smirnov test, with p<0,001 (***) for sympatry. **b.** Frequency of 4-spore tetrads in homozygous and heterozygous crosses throughout time in population B8. Error bars are twice the standard error (2SE) of 20 to 40 crosses per timepoint. Statistical significance was inferred with the Kolmogorov-Smirnov test, with * p<0,05, ** p<0,01, and *** p<0,001. **c.** Meiotic viability assessed by random spore analysis (RSA), representing the results for at least 7 independent experiments replicated 6 times each in population B8. Error bars are 2SE, and a two-tailed T-test was used to compare the heterozygous ancestral and evolved viabilities with p>0,05 (ns).

IV.4. Whole-chromosome recombination rates remain unchanged after 500 generations of sympatric sexual evolution

We wanted to assess whether sympatry had led to selection of suppression of recombination. We hypothesized that the increase in production of spores containing the correct amount of DNA was due to evolution of suppression of recombination in chromosome I in the heterozygous meiotic pairing. To test this, we constructed strains containing a marker

halfway to the breakpoints (in chromosome I) from the evolved and ancestral clones and measured recombination rates in heterozygous and homozygous crosses. For this assay, we used the isolated clones from population B8 where we observed the increased healthy tetrad production – a WT h+ and a Inv h- – as well as the ancestral strains of equivalent genotypes.

We found that in ancestral homozygous crosses, as expected, the recombination rate is roughly 50%. This is because there would be no problems in alignment as the homologue chromosomes are collinear, and thus recombine freely. In ancestral heterozygous crosses, the recombination rate in chromosome I is 30%, which is surprisingly high. In evolved heterozygous crosses, the recombination rate is the same as in the ancestral, 30% (figure 9). Thus we observed no change in the recombination rates after 500 generations of evolution and can exclude it as a cause for the increased production of balanced 4-spore tetrads.

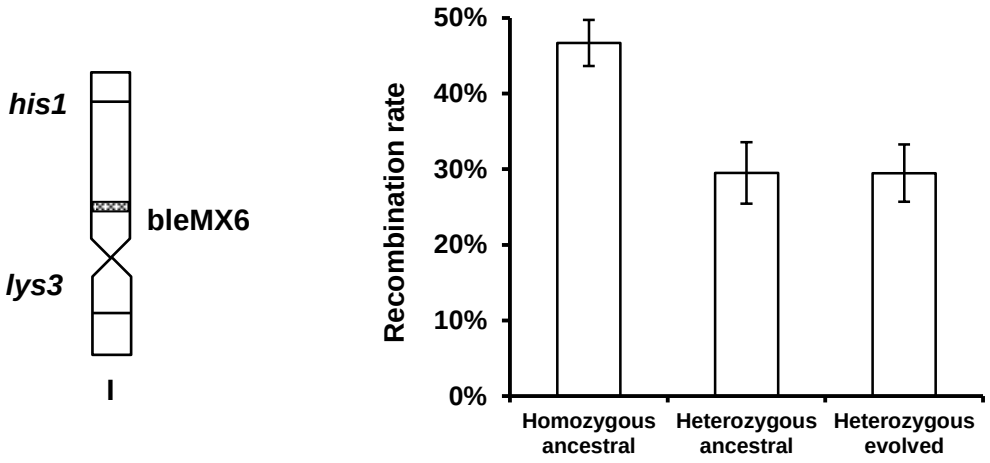


Figure 9. Recombination rates in homozygous and heterozygous crosses. The measured values refer to the recombination rate found between the cassette bleMX6 and the *his1* breakpoint where the fluorescent marker is located. Because breakpoints are linked, the rate is the same for the *lys3* breakpoint. Graphs represent the average of 9-14 independent experiments, and error bars are 2SE of the mean. The diagram shows the location of the marker used to measure recombination relatively to the breakpoints in chromosome I. The marker was inserted in the WT background which was used to cross with either other WT cells or Inv cells.

V. Discussion

Chromosomal rearrangements are an obstacle to correct recombination, as they can lead to structurally unviable products (figure 2). Surviving offspring will carry the breakpoints in linkage, and, according to the hypothesis, accumulate mutations nearby leading to karyotypic differentiation. Suppression of recombination in the presence of epistasis could arise in the case of sporadic allele exchange[44]. Alternatively, there could be selection for other mechanisms of coping with the problem, such as new rearrangements, or increasing recombination in order to increase the probability of balanced crossover products. We have devised an evolution experiment to test the chromosomal speciation hypothesis. For 500 generations, fission yeast cells with an inversion have evolved with WT cells in long periods of mitosis intercalated with rounds of meiosis. One of the aims of this project was to verify if selection can establish a population of individuals who have suppressed recombination, thus reducing the gene flow between strains to avoid meiotic fitness depression, or as a result of epistasis between divergently accumulated mutations in linkage with breakpoints. This reduction of gene flow could eventually lead to chromosomal dimorphism and ultimately speciation. We aimed to find out which molecular and genetic mechanisms can arise in sexual populations posed with strong recombination problems.

Our main observation was the evolution of sterility as an escape mechanism in at least half of the sexual populations evolved in sympatric conditions. More than 30% of the sympatric sexual populations fixed sterility, while more than 20% were composed of sterile cells coexisting with mating-proficient cells. It is likely we are underestimating the real number of populations where sterility arose, because we did not exhaustively estimate the frequency of sterile clones coexisting with others in populations where a single chromosome structure was fixed (table 1). Whole genome analysis of two sterile clones from independent populations showed no mutations in common between them (figure S4). This suggests that sterility was the selected phenotype, and not the result of a trade-off or a hitchhiking mutation. Additionally, we found that sterility appeared independently in population B8 at several points in the experiment. The first sterile clone observed, a WT, was prior to the third round of meiosis. After meiosis, sterility appeared again in an Inv background. Because of the phenotype it confers, the allele cannot be transmitted through mating. The occurrence of multiple events of evolution of sterility is the simplest explanation for the appearance of the trait in different chromosome configuration backgrounds. These observations, coupled with the fact that the trait first appeared early in the experiment, suggest that selection for it was strong and the mutational target was large. We also observed that sterility did not appear in sexual allopatric or in asexual sympatric circumstances. Hence, strong selection for sterility

was a direct consequence of sexual selection and co-existence of two chromosome structures.

Analysis of the sexual sympatric populations' evolutionary trajectories (figure 5a) revealed a common phenomenon to all, the post-meiosis increase in frequency of the inversion. By exploring this, we observed that in sympatry fitness differences prior to meiosis, by influencing genotype frequencies, can have direct consequences on the populations' fates. In this case, they indirectly catalysed the selection for escaping sex. In our experiment, the WT h+ increased very quickly in frequency before the first meiosis, while the WT h- was almost lost by generation 100 (figure 5c and asexual data, also happened in other polymorphic populations such as B6). This meant that WT cells had only one mating option, Inv cells, while these retained both options (at least in population B8, figure 5c). Because WT h+ was at a high frequency, virtually all crosses that occurred in the first round of meiosis were heterozygous, as confirmed by tracing a WT h+ specific allele. We found that heterozygous crosses between WT and Inv chromosomes have 60% of lethality (figure 8), which in itself translates into a high selection coefficient for mutations that resolve the structural incompatibility, such as sterility. The same scenario was verified in the second round of meiosis, given the relative frequencies of the four genotypes at this point (figure 5c). High hybrid lethality, coupled with the frequency of heterozygous crosses at both these points, likely resulted in strong selection for coping mechanisms.

Sterile cells may also have had an additional advantage, due to the nature of the experiment. Selection of sexual products by snail juice usage likely contributed to the appearance of sterile cells because cells which are able to survive the process without undergoing meiosis not only suffer no meiotic depression but also grow faster than the spores germinate, occupying the niche quicker. Hence a sterile strain, able to survive the snail juice treatment should be highly selected for. This is corroborated by the observation that sterility is significantly correlated with resistance to snail juice (figure 7).

It has been previously described that strong sexual selection is detrimental to populations because it delays adaptation[63]. The authors of this study hypothesize that strong sexual selection may favour alleles that facilitate sex, but that, by epistasis or otherwise, this may come with a fitness cost and delay adaptation. If there is such a trade-off, we expect sterile cells to be more well adapted to the environment than mating-proficient ones. In order to test whether this is the case, we should compare the fitness of the sterile and non-sterile evolved strains. We expect the sterile cells to have an advantage both in the mitotic growth and during the snail juice treatment.

We found that meiosis is crucial in maintaining chromosome structure diversity (figure 5a), by resetting the frequencies of each genotype following heterozygous crosses (figure 5d shows the increase of the Inv background to roughly 50% following the first round of meiosis,

which, as mentioned, consisted mainly in heterozygous crosses). This is compatible with previous observations from Avelar et. al, 2013[6]: meiosis counterbalances mitotic advantages by antagonistic pleiotropy, delaying fixation of beneficial variants. By eliminating sex and therefore recombination, sterility uncoupled the fates of each structure genotype from each other. Thus, after the selection of the sterile phenotypes, the evolution of chromosome structure is expected to be dominated by mitotic competition. Beneficial structures and hitchhiking alleles will fixate faster, increasing divergence and decreasing variability. This is compatible with our observations. We observed that populations where sterility is fixed not only have lost mating type polymorphism (table 1), but have also fixed one chromosome structure (mainly WT, with only B12 fixing the inversion – table 1). If sterility was fixed early in their evolution, then clonal interference may have been responsible for this outcome after 500 generations, with beneficial mutations appearing in one background, being trapped in it without meiosis, and taking over the population. Additionally, in the long-term, sterility could lead to divergence between populations, because it also eliminates the possibility of negative epistasis arising between mutations come from different structural backgrounds which would be selected against.

In one particular population, B8, we observed a phenotype of increased production of balanced, healthy asci resulting from heterozygous crosses that was not accompanied by an expected increase in overall spore viability (figure 8), which we cannot explain. We hypothesized that by balancing DNA content in the spores after meiosis, the probability of products being viable afterwards is increased, but such an increase was not observed in our assays. Furthermore, tetrad dissection results showed that these 4-spore asci maintained the ancestral low viability of heterozygous crosses (roughly 50%, figure S3). However, this phenotype seems to have been selected throughout the experiment, from the third meiosis on (figure 8c), possibly because the appearance of sterility in the Inv h+ background influenced the amount of mating partners available for the Inv h- cells. We have verified that this trait is not a consequence of changing recombination rates in the evolved clones (figure 9). Whole-genome sequencing of clones presenting this phenotype have revealed new mutations appearing during the experiment (figure S4), but we could not attribute the phenotype to any of the mutations, and so the mechanism by which it manifests remains elusive.

As mentioned, we did not observe a change in recombination rates in chromosome I in this experiment. Indeed, we measured a rate of 30% of recombination (between breakpoints and mid-chromosome marker) appearing in heterozygous incompatible crosses, both ancestral and evolved. This is an indication that the phenomenon of crossing-over continues to occur quasi-normally, as at least one crossover at each side of the marker is sufficient to produce a balanced recombinant (because distance between breakpoints and marker is over

1Mb). This is compatible with the hypothesis that a balanced number of crossovers in a heterozygous pairing can prevent lethality. However, an odd number of crossovers leads to lethality, which would explain the decrease of 50% to 30% of recombination rate of homozygous to heterozygous crosses, respectively. These data differ from our data for chromosome II[6]: for heterozygous matings with an inversion in chromosome II, the recombination rate is lower than 20%. Chromosome size (I: 5,6Mb vs. II: 4,5Mb) could account for a percentage of the difference, but it is unlikely to fully explain it. We did not measure recombination rates closer to the breakpoints; nonetheless, we observed accumulation of mutations near the breakpoints in evolved cells that were not shared between structures (figure S4, refer to clones AT395 and AT387 for example). We have consistently verified complete linkage between breakpoints in this experiment, as previously[6].

In this experiment, we found no evidence for further decrease of meiotic viability indicating post-zygotic isolation, or for suppression of recombination. Instead, cells quickly escaped the sexual selection process by becoming sterile. Because of this, we were not able to directly test the chromosome speciation hypothesis. We were, however, able to verify meiotic depression between incompatible genomes and we observed that mutations can indeed accumulate near the breakpoints in this system (figure S4). Ideally, the hypothesis should be formally tested in an organism that can only reproduce sexually, in order to not only increase selection but also guarantee there is no escape from meiosis. In summary, the chromosome speciation hypothesis remains untested. Instead, we observed the unexpected evolution of a trait that could also theoretically lead to separation.

The emergence of asexual lineages can lead to very rapid strain divergence. In this experiment, we found a new mechanism that may lead to genetic divergence between different karyotypes in facultative sexual populations. Butlin, 2005[64] defined speciation as the evolution of restriction on the freedom of genetic recombination. Some sympatric speciation models[64,65] state that speciation occurs when restricted recombination is favoured by selection and that its suppression enhances progress towards separation. Although we did not find evidence for restricted recombination in mating-proficient cells in our experimental timeframe, we found complete and abrupt recombination abolition in a subpopulation of cells which were favoured by selection due to genome incompatibility.

VI. References

1. Dekker, J. The three “C” s of chromosome conformation capture: controls, controls, controls. *Nat. Methods* **3,** 17–21 (2006).
2. Van Berkum, N. L. *et al.* Hi-C: a method to study the three-dimensional architecture of genomes. *J. Vis. Exp.* (2010). doi:10.3791/1869
3. Zhang, H., Liu, D.-P. & Liang, C.-C. The control of expression of the alpha-globin gene cluster. *Int. J. Hematol.* **76,** 420–426 (2002).
4. Querol, A. Adaptive evolution of wine yeast. *Int. J. Food Microbiol.* **86,** 3–10 (2003).

5. Dobzhansky, T. & Sturtevant, A. Inversions in the Chromosomes of Drosophila Pseudoobscura. *Genetics* (1938).
6. Avelar, A. T., Perfeito, L., Gordo, I. & Ferreira, M. G. Genome architecture is a selectable trait that can be maintained by antagonistic pleiotropy. *Nat. Commun.* **4,** 2235 (2013).
7. Anderson, W. W. *et al.* Four decades of inversion polymorphism in Drosophila pseudoobscura. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 10367–10371 (1991).
8. Dunham, M. J. *et al.* Characteristic genome rearrangements in experimental evolution of Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 16144–16149 (2002).
9. Albertson, D. G., Collins, C., McCormick, F. & Gray, J. W. Chromosome aberrations in solid tumors. *Nat. Genet.* **34,** 369–376 (2003).
10. Nambiar, M., Kari, V. & Raghavan, S. C. Chromosomal translocations in cancer. *Biochim. Biophys. Acta - Rev. Cancer* **1786,** 139–152 (2008).
11. Nordling, C. O. A new theory on cancer-inducing mechanism. *Br. J. Cancer* **7,** 68–72 (1953).
12. Knudson, A. G. Two genetic hits (more or less) to cancer. *Nat. Rev. Cancer* **1,** 157–162 (2001).
13. Stephens, P. J. *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144,** 27–40 (2011).
14. Liu, B. *et al.* Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. *Nat. Med.* **2,** 169–174 (1996).
15. Willis, N. a *et al.* BRCA1 controls homologous recombination at Tus/Ter-stalled mammalian replication forks. *Nature* **510,** 556–9 (2014).
16. Aguilera, A. & Gómez-González, B. Genome instability: a mechanistic view of its causes and consequences. *Nat. Rev. Genet.* **9,** 204–217 (2008).
17. Van Gent, D. C., Hoeijmakers, J. H. & Kanaar, R. Chromosomal stability and the DNA double-stranded break connection. *Nat. Rev. Genet.* **2,** 196–206 (2001).
18. Gu, W., Zhang, F. & Lupski, J. R. Mechanisms for human genomic rearrangements. *Pathogenetics* **1,** 4 (2008).
19. Kasparek, T. R. & Humphrey, T. C. DNA double-strand break repair pathways, chromosomal rearrangements and cancer. *Semin. Cell Dev. Biol.* **22,** 886–897 (2011).
20. Reis, C. C., Batista, S. & Ferreira, M. G. The fission yeast MRN complex tethers dysfunctional telomeres for NHEJ repair. *EMBO J.* **31,** 4576–86 (2012).
21. Weterings, E. & Chen, D. J. The endless tale of non-homologous end-joining. *Cell Res.* **18,** 114–124 (2008).
22. Hurles, M. E. & Lupski, J. R. in *Genomic Disord. Genomic Basis Dis.* 341–355 (2006). doi:10.1007/978-1-59745-039-3_24
23. Hastings, P. J., Lupski, J. R., Rosenberg, S. M. & Ira, G. Mechanisms of change in gene copy number. *Nat. Rev. Genet.* **10,** 551–564 (2009).
24. Ranz, J. M. *et al.* Principles of genome evolution in the Drosophila melanogaster species group. *PLoS Biol.* **5,** 1366–1381 (2007).
25. Bickhart, D. M. & Liu, G. E. The challenges and importance of structural variation detection in livestock. *Front. Genet.* **5,** 37 (2014).
26. Bunting, S. F. & Nussenzweig, A. End-joining, translocations and cancer. *Nat. Rev. Cancer* **13,** 443–54 (2013).
27. Kehrer-Sawatzki, H., Sandig, C. A., Goidts, V. & Hameister, H. Breakpoint analysis of the pericentric inversion between chimpanzee chromosome 10 and the homologous chromosome 12 in humans. *Cytogenet. Genome Res.* **108,** 91–97 (2005).
28. McClintock, B. The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics* **26,** 234–282 (1941).
29. Hu, J., Tepsuporn, S., Meyers, R. M., Gostissa, M. & Alt, F. W. Developmental propagation of V(D)J recombination-associated DNA breaks and translocations in mature B cells via dicentric chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 10269–10274 (2014).
30. Murnane, J. P. Telomere dysfunction and chromosome instability. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **730,** 28–36 (2012).
31. Nowell, P. C. A minute chromosome in human chronic granulocytic leukemia. *Science (80-.).* **8,** 19 (1985).
32. Deininger, M. W. N., Goldman, J. M. & Melo, J. V. The molecular biology of chronic myeloid leukemia. *Blood* **96,** 3343–3356 (2000).
33. Adams, J. M., Gerondakis, S., Webb, E., Corcoran, L. M. & Cory, S. Cellular myc oncogene is altered by chromosome translocation to an immunoglobulin locus in murine plasmacytomas and is rearranged similarly in human Burkitt lymphomas. *Proc. Natl. Acad. Sci. U. S. A.* **80,** 1982–1986 (1983).

34. Owen, L. a, Kowalewski, A. a & Lessnick, S. L. EWS/FLI mediates transcriptional repression via NKX2.2 during oncogenic transformation in Ewing's sarcoma. *PLoS One* **3,** e1965 (2008).
35. Anderson, A. R., Hoffmann, A. A., Mckechnie, S. W., Umina, P. A. & Weeks, A. R. The latitudinal cline in the In(3R)Payne inversion polymorphism has shifted in the last 20 years in Australian Drosophila melanogaster populations. *Mol. Ecol.* **14,** 851–858 (2005).
36. Lowry, D. B. & Willis, J. H. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **8,** (2010).
37. Dunn, B. *et al.* Recurrent Rearrangement during Adaptive Evolution in an Interspecific Yeast Hybrid Suggests a Model for Rapid Introgression. *PLoS Genet.* **9,** e1003366 (2013).
38. Avelar, A. T. Chromosomal structure : a selectable trait for evolution. (2012).
39. Loidl, J., Jin, Q. W. & Jantsch, M. Meiotic pairing and segregation of translocation quadrivalents in yeast. *Chromosoma* **107,** 247–254 (1998).
40. Brown, W. R. A. *et al.* A Geographically Diverse Collection of Schizosaccharomyces pombe Isolates Shows Limited Phenotypic Variation but Extensive Karyotypic Diversity. *G3 Genes|Genomes|Genetics* **1,** 615–626 (2011).
41. Ayala, F. J. & Coluzzi, M. Chromosome speciation: humans, Drosophila, and mosquitoes. *Proc. Natl. Acad. Sci. U. S. A.* **102 Suppl ,** 6535–6542 (2005).
42. White, M. J. D. Models of Speciation. *Science (80-.).* **8,** 1065–1070 (1968).
43. Rieseberg, L. H. Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* **16,** 351–358 (2001).
44. Faria, R. & Navarro, A. Chromosomal speciation revisited: Rearranging theory with pieces of evidence. *Trends Ecol. Evol.* **25,** 660–669 (2010).
45. Navarro, A., Betrán, E., Barbadilla, A. & Ruiz, A. Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* **146,** 695–709 (1997).
46. Coyne, J. A., Aulard, S. & Berry, A. Lack of Underdominance in a Naturally-Occurring Pericentric-Inversion in Drosophila-Melanogaster and Its Implications for Chromosome Evolution. *Genetics* **129,** 791–802 (1991).
47. Noor, M. A., Grams, K. L., Bertucci, L. A. & Reiland, J. Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 12084–12088 (2001).
48. Bryan, J. H., Deco, M. A., Petrarca, V. & Coluzzi, M. Inversion polymorphism and incipient speciation in Anopheles gambiae s.str. in The Gambia, West Africa. *Genetica* **59,** 167–176 (1982).
49. Kirkpatrick, M. & Barton, N. Chromosome inversions, local adaptation and speciation. *Genetics* **173,** 419–434 (2006).
50. Dobzhansky, T. *Genetics and the origin of species*. *New York Columbia Univ. Press* (1937). at <http://www.elib.gov.ph/details.php?cat_id=429494>
51. Navarro, A. & Barton, N. H. Chromosomal speciation and molecular divergence--accelerated evolution in rearranged chromosomes. *Science* **300,** 321–324 (2003).
52. Navarro, A. & Barton, N. H. Accumulating Postzygotic Isolation Genes in Parapatry: A New Twist on Chromosomal Speciation. *Evolution (N. Y).* **57,** 447–459 (2003).
53. Rieseberg, L. H., Van Fossen, C. & Desrochers, A. M. Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* **375,** 313–316 (1995).
54. Rieseberg, L. H., Whitton, J. & Gardner, K. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* **152,** 713–727 (1999).
55. Kirkpatrick, M. How and why chromosome inversions evolve. *PLoS Biol.* **8,** (2010).
56. Lahn, B. T. & Page, D. C. Four evolutionary strata on the human X chromosome. *Science* **286,** 964–967 (1999).
57. Forsburg, S. L. & Rhind, N. Basic methods for fission yeast. *Yeast* **23,** 173–83 (2006).
58. Moreno, S., Klar, A. & Nurse, P. Molecular Genetic Analysis of Fission Yeast Schizosaccharomyces pombe. *Methods Enzymol.* **194,** 795–823 (1991).
59. Tomita, K. & Cooper, J. P. The Telomere Bouquet Controls the Meiotic Spindle. *Cell* **130,** 113–126 (2007).
60. Hoffman, C. S. & Winston, F. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli. *Gene* **57,** 267–272 (1987).
61. Hentges, P., Van Driessche, B., Tafforeau, L., Vandenhaute, J. & Carr, A. M. Three novel antibiotic marker cassettes for gene disruption and marker switching in Schizosaccharomyces pombe. *Yeast* **22,** 1013–1019 (2005).
62. Milne, I. *et al.* Using tablet for visual exploration of second-generation sequencing data. *Brief. Bioinform.* **14,** 193–202 (2013).

63. Reding, L. P., Swaddle, J. P. & Murphy, H. A. Sexual selection hinders adaptation in experimental populations of yeast. *Biol. Lett.* **9,** 9–13 (2013).
64. Butlin, R. K. Recombination and speciation. *Mol. Ecol.* **14,** 2621–2635 (2005).
65. Faria, R., Neto, S., Noor, M. & Navarro, A. in *Encycl. Life Sci.* 1–14 (John Wiley & Sons, Ltd, 2011). doi:10.1002/9780470015902.a0022850

Appendices

Supplementary figures

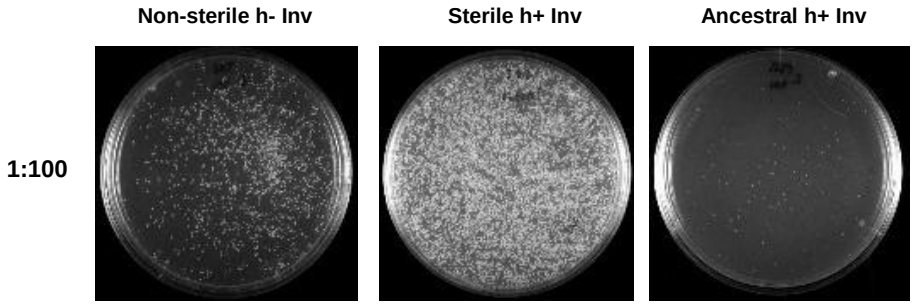


Figure S1. Sterility correlates with SJ resistance. Representative plates of several clones after SJ treatment at 1:100, with about 10^5 cells plated.

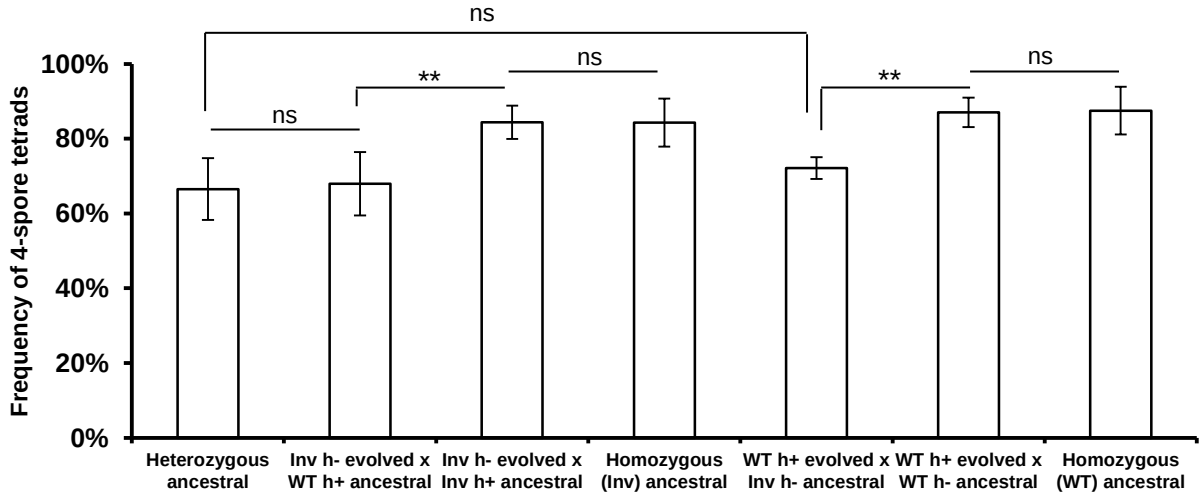


Figure S2. 4-spore tetrad proportion for evolved cells crosses with ancestral cells. Values refer to the mean of 4-6 repetitions of each cross. Ancestral crosses' values are the mean of the values used for the distributions in figure 8a. Error bars are 2SE. Statistical significance was assessed using a two-tailed T-test, with p<0,01 (**) and p>0,05 (ns).

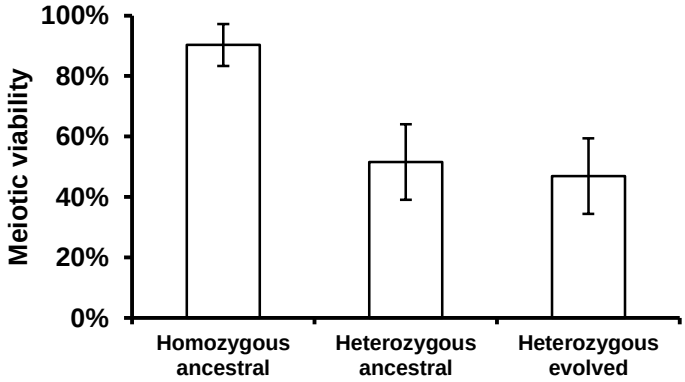
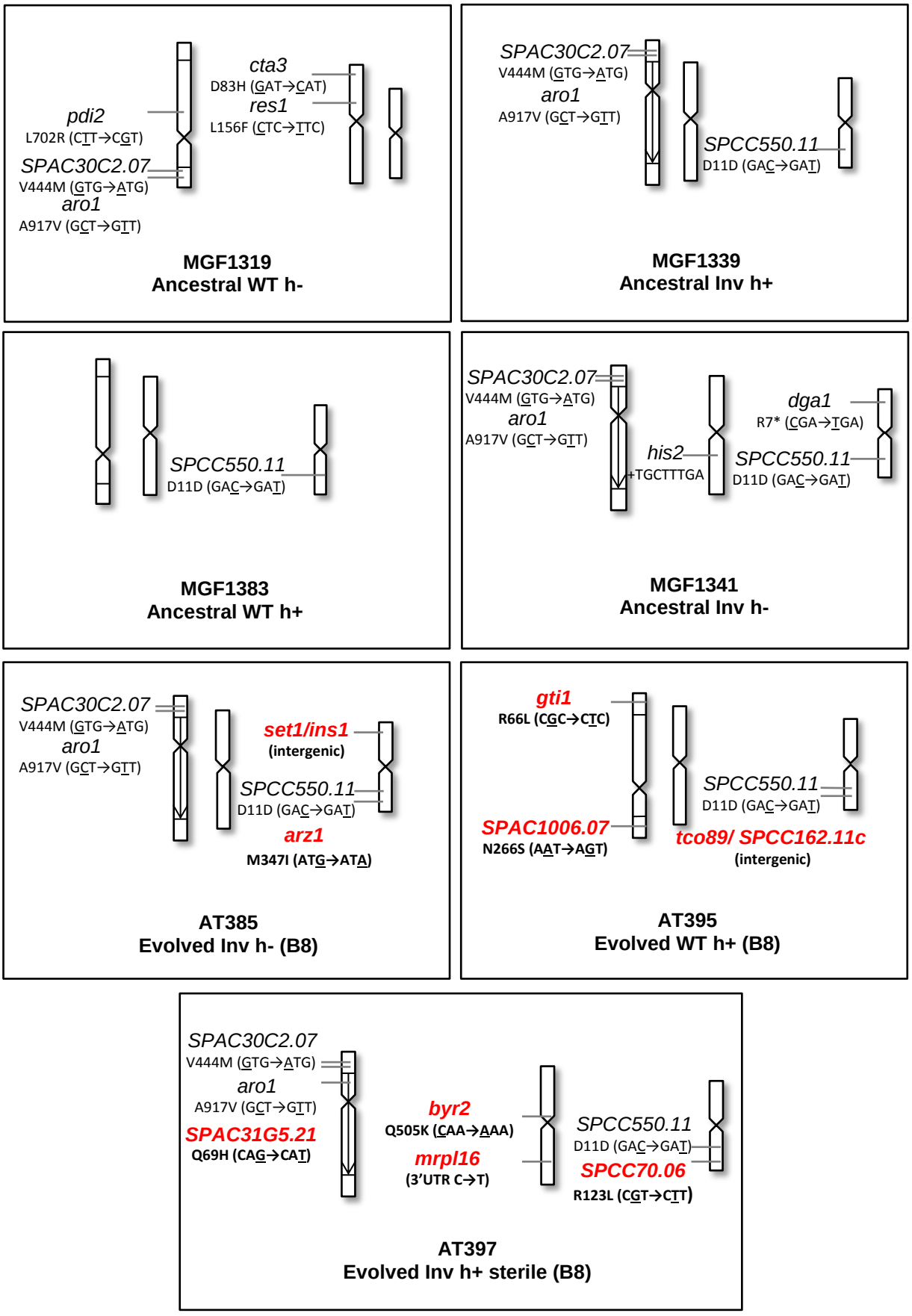


Figure S3. Meiotic viabilities by tetrad dissection. Between 16 and 20 tetrads were dissected for each cross. Graphs represent the average of 3 independent experiments. Error bars 2SE of a binomial distribution.


pdi2
L702R (CTT→CGT)
SPAC30C2.07
V444M (GTG→ATG)
aro1
A917V (GCT→GTT)
cta3
D83H (GAT→CAT)
res1
L156F (CTC→TTC)
MGF1319
Ancestral WT h-
SPAC30C2.07
V444M (GTG→ATG)
aro1
A917V (GCT→GTT)
SPCC550.11
D11D (GAC→GAT)
MGF1339
Ancestral Inv h+
SPCC550.11
D11D (GAC→GAT)
MGF1383
Ancestral WT h+
SPAC30C2.07
V444M (GTG→ATG)
aro1
A917V (GCT→GTT)
his2
+TGCTTTGA
dga1
R7* (CGA→TGA)
SPCC550.11
D11D (GAC→GAT)
MGF1341
Ancestral Inv h-
SPAC30C2.07
V444M (GTG→ATG)
aro1
A917V (GCT→GTT)
set1/ins1
(intergenic)
SPCC550.11
D11D (GAC→GAT)
arz1
M347I (ATG→ATA)
AT385
Evolved Inv h- (B8)
gti1
R66L (CGC→CTC)
SPAC1006.07
N266S (AAT→AGT)
SPCC550.11
D11D (GAC→GAT)
tco89/ SPCC162.11c
(intergenic)
AT395
Evolved WT h+ (B8)
SPAC30C2.07
V444M (GTG→ATG)
aro1
A917V (GCT→GTT)
SPAC31G5.21
Q69H (CAG→CAT)
byr2
Q505K (CAA→AAA)
mrpl16
(3'UTR C→T)
SPCC550.11
D11D (GAC→GAT)
SPCC70.06
R123L (CGT→CTT)
AT397
Evolved Inv h+ sterile (B8)

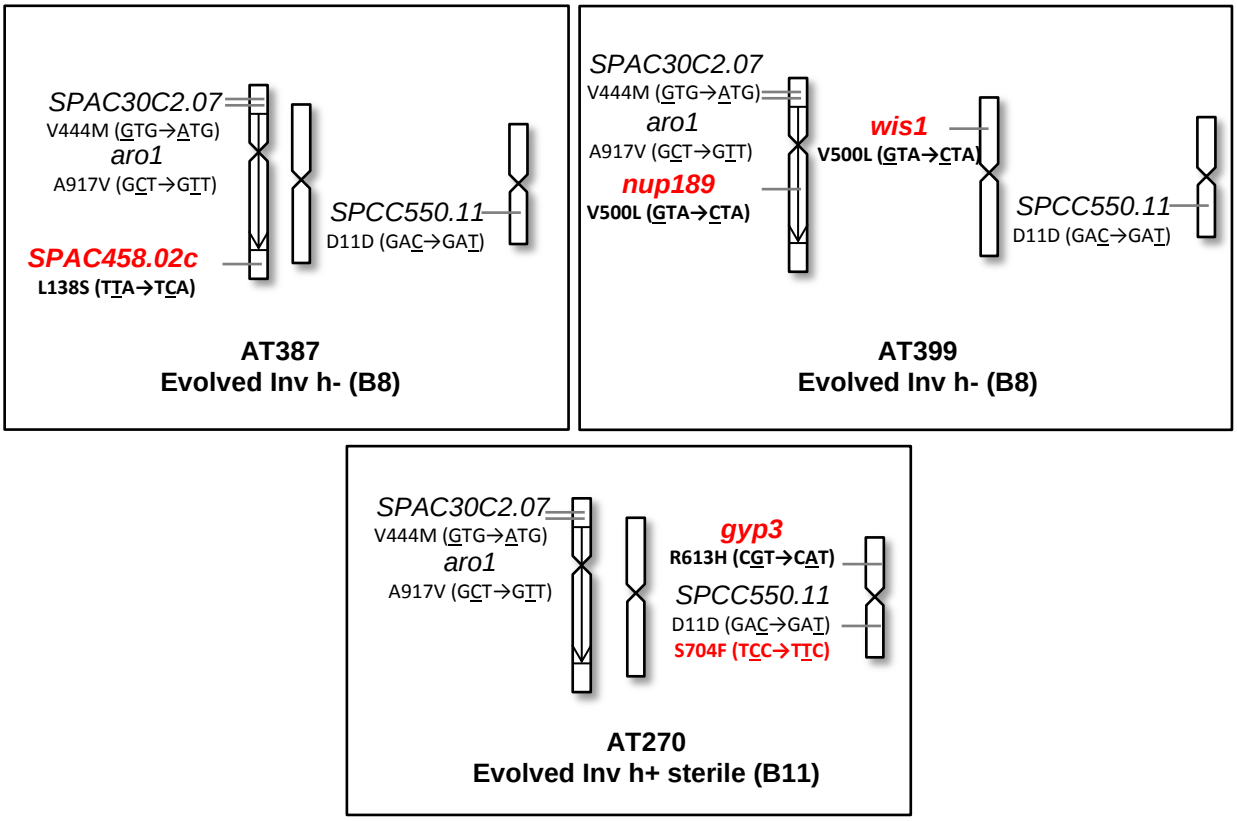


Figure S4. Mutations found in this study and their locations in the genome. These mutations were assigned relatively to the reference strain. In bold red, new mutations of evolved clones relatively to the ancestral ones. All mutations are non-synonymous, except the mutation in SPCC550.11.

Supplementary tables

Table S1. List of strains used in this study. n.a. = not applicable.

Strain name	Genotype	Common name	Creator
MGF10	*h- ade6-M210 his3-D1 leu1-32 ura4-D18*	n.a.	Richard McIntosh
MGF11	*h+ ade6-M210 his3-D1 leu1-32 ura4-D18*	n.a.	Richard McIntosh
MGF846	*h- L972 matM:natMX6*	n.a.	Avelar et al.[6]
MGF847	*h+ L972 matP:natMX6*	n.a.	Avelar et al.[6]
MGF1319	*h- lys3::padh1-loxP-kanMX6 his1::loxP-ura4-kanMX6- mCherry-hphMX6 mat1-M leu1-32 pdi2-L702R SPAC30C2.07-V444M aro1-A917V cta3-D83M res1-C156F*	Ancestral control h-	Avelar et al.[6]
MGF1341	*h- his1::loxP-kanMX6-GFP-hphMX6 lys3::padh1-loxP-ura4-kanMX6 mat1-M::mat1-M-natMX6 leu1-32 ade6-M216 ura4-D18 dga1-R7* SPCC550.11-D11D his2-c.428_435insTGCTTTGA*	Ancestral inverted h-	Avelar et al.[6]

MGF1339	*h+ his1::loxP-kanMX6-GFP-hphMX6 lys3::padh1-loxP-ura4-kanMX6 mat1-P::mat1-P-natMX6 leu1-32 ade6-M216 ura4-D18 SPAC30C2.07-V444M aro1-A917V SPCC550.11-D11D*	Ancestral inverted h+	Avelar et al.[6]
MGF1383	*h+ lys3::padh1-loxP-kanMX6 his1::loxP-ura4-kanMX6- mCherry-hphMX6 mat1-P::mat1-P-natMX6 ade6-M216 SPCC550.11-D11D*	Ancestral control h+	Avelar et al.[6]
AT385	*h- his1::loxP-kanMX6-GFP-hphMX6 lys3::padh1-loxP-ura4-kanMX6 mat1-M::mat1-M-natMX6 ade6-M216 ura4-D18 arz1-M347I SPAC30C2.07-V444M aro1-A917V SPCC550.11-D11D ins1-c.+107G>A*	Evolved inverted h- (population B8)	This study
AT395	*h+ lys3::padh1-loxP-kanMX6 his1::loxP-ura4-kanMX6- mCherry-hphMX6 mat1-P::mat1-P-natMX6 ade6-M216 gti1-R66L SPAC1006.07-N266S SPCC550.11-D11D tco89-c.-1095T>G*	Evolved control h+ (population B8)	This study
AT397	*h+ his1::loxP-kanMX6-GFP-hphMX6 lys3::padh1-loxP-ura4-kanMX6 mat1-P::mat1-P-natMX6 ade6-M216 ura4-D18 SPCC70.06-R123L byr2-Q505K SPAC31G5.21-Q69H SPAC30C2.07-V444M aro1-A917V SPCC550.11-D11D mrpl16-c.+745C>T*	Evolved inverted h+ sterile (population B8)	This study
AT387	*h- his1::loxP-kanMX6-GFP-hphMX6 lys3::padh1-loxP-ura4-kanMX6 mat1-M::mat1-M-natMX6 ade6-M216 ura4-D18 SPAC458.02c-L138S SPAC30C2.07-V444M aro1-A917V SPCC550.11-D11D*	Evolved inverted h- (population B8)	This study
AT399	*h- his1::loxP-kanMX6-GFP-hphMX6 lys3::padh1-loxP-ura4-kanMX6 mat1-M::mat1-M-natMX6 ade6-M216 ura4-D18 wis1-V500L nup189-K82R SPAC30C2.07-V444M aro1-A917V SPCC550.11-D11D*	Evolved inverted h- (population B8)	This study
AT270	*h+ his1::loxP-kanMX6-GFP-hphMX6 lys3::padh1-*	Evolved inverted h+ sterile (population B11)	This study

	loxP-ura4-kanMX6 mat1-P::mat1-P-natMX6 ade6-M216 ura4-D18 gyp3-R613H SPAC30C2.07-V444M aro1-A917V SPCC550.11-D11D SPCC50.11-S704F		
MGF2500	*h+ lys3::padh1-loxP-kanMX6 his1::loxP-ura4-kanMX6- mCherry-hphMX6 mat1-P::mat1-P-natMX6 ade6-M216 gti1-R66L SPAC1006.07-N266S SPCC550.11-D11D tco89-c.-1095T>G I:2600293::bleMX6*	AT395 bleMX6^{R}	This study
MGF2499	*h+ lys3::padh1-loxP-kanMX6 his1::loxP-ura4-kanMX6- mCherry-hphMX6 mat1-P::mat1-P-natMX6 ade6-M216 SPCC550.11-D11D I:2600293::bleMX6*	MGF1383 bleMX6^{R}	This study
MGF2501	*h- lys3::padh1-loxP-kanMX6 his1::loxP-ura4-kanMX6-GFP-hphMX6 leu1-32 ura4+ ade6-M216 mat1-M::mat1-M-natMX6R*	Control GFP	This study

Table S2. List of primers used in this study.

Primer name	Sequence 5'-3'	For
MM	acgttcagtagacgtagtg	Mating type amplification
MP	acggtagtcatcggtcttcc	Mating type amplification
MT1	agaagagagagtagttgaag	Mating type amplification
his1-100 forw	gactgctttttcgacatgg	Breakpoint amplification
lys3-200 forw	agtttttggtctccttcgcc	Breakpoint amplification
kanMX6-880 rev	cgcatcaaccaaaccgttat	Breakpoint amplification
F2	cggatccccgggttaattaa	MX6 cassette amplification
R1	gaattcgagctcgtttaaac	MX6 cassette amplification
A-forward	gcttcttcgcagttgtgtaagc	MX6 cassette insertion in chromosome I
B-reverse	ttaattaacccggggatccgctcattgaacttgagatagtg	MX6 cassette insertion in chromosome I
C-forward	gtttaaacgagctcgaattccccgaacttaacagttgagc	MX6 cassette insertion in chromosome I
D-reverse	cgcaattacatctgaagctg	MX6 cassette insertion in chromosome I
check bleMX6 F	ctgaaaggacaaccaggaac	Insertion confirmation
check bleMX6 R	ggtggtgaaggttgaaatgc	Insertion confirmation

Supplementary notes

Post-meiosis frequency change model. A model was developed in order to predict frequency changes of the wild type and inverted genotypes after each round of meiosis, based on the frequencies measured at points during the mitotic phases. The model is given by

$$F_{I(t+1)} = \frac{4F_{I+(t)}F_{I-(t)} + F_{I+(t)}F_{WT-(t)} + F_{I-(t)}F_{WT+(t)}}{4F_{I+(t)}F_{I-(t)} + 4F_{WT+(t)}F_{WT-(t)} + 2F_{I+(t)}F_{WT-(t)} + 2F_{I-(t)}F_{WT+(t)}} \quad (1)$$

where the frequencies for time t are considered as the values measured experimentally by clone sampling. The numerator represents the absolute frequency of cells containing the inverted genotype after meiosis, considering that I. the probability of a sexual encounter between two cells is given by their frequencies multiplied, II. crosses between incompatible genotypes produce progeny consisting of 50% inverted and 50% wild type cells, and III. crosses between incompatible genomes yield a 50% lethality rate. The proportion, or relative frequency, of inverted cells in the population is thus given by dividing the total frequency by the total number of alive cells post-meiosis. At time t+1 the frequency of wild type cells is given by

$$F_{WT(t+1)} = 1 - F_{I(t+1)} \quad (2)$$

The model was then adjusted by correcting the sexual encounter probabilities for measured meiotic viabilities of ancestral crosses (90% for homozygous crosses and 40% for heterozygous – see figure 8c), thus becoming

$$F_{I(t+1)} = \frac{4{,}5F_{I+(t)}F_{I-(t)} + F_{I+(t)}F_{WT-(t)} + F_{I-(t)}F_{WT+(t)}}{4{,}5F_{I+(t)}F_{I-(t)} + 4{,}5F_{WT+(t)}F_{WT-(t)} + 2F_{I+(t)}F_{WT-(t)} + 2F_{I-(t)}F_{WT+(t)}} \quad (3)$$

The model does not calculate post-meiosis mating type frequencies. It assumes that segregation of the *mat1* locus is independent of the inversion. This was verified experimentally. As such, the 50/50 proportion of h+/h- cells assumed in the frequency values given by applying the model to experimental data is not comparable to the actual values measured by clone sampling. In addition to the intrinsic error of the sampling process, these points were taken after a period of post-meiosis mitotic propagation, which in itself alters the 50/50 proportion.