Published in final edited form as:
Nature. 2021 June 01; 594(7864): 572–576. doi:10.1038/s41586-021-03389-3.

Concerted cutting by Spo11 illuminates meiotic DNA break mechanics

Dominic Johnson[#1], **Margaret Crawford**[#1], **Tim Cooper**[1], **Corentin Claeys Bouuaert**[3,4], **Scott Keeney**[3], **Bertrand Llorente**[5], **Valerie Garcia**[1,5,6], **Matthew J. Neale**[1,6]

[1]Genome Damage and Stability Centre, University of Sussex, Brighton, UK

[3]Howard Hughes Medical Institute, Memorial Sloan Kettering Cancer Center, New York, USA

[4]Louvain Institute of Biomolecular Science and Technology, Louvain-la-Neuve, Belgium

[5]Cancer Research Centre of Marseille, CNRS, Inserm, Institut Paoli-Calmettes, Aix-Marseille Université, Marseille, France

[#] These authors contributed equally to this work.

Abstract

Genetic recombination arises during meiosis via repair of DNA double-strand breaks (DSBs) created by the topoisomerase-like Spo11 protein[1,2]. Spo11 DSBs form preferentially in nucleosome-depleted regions termed hotspots[3,4], yet how Spo11 engages with its DNA substrate to catalyse DNA cleavage is poorly understood. Whilst most recombination events are initiated by a singular Spo11 cut, we demonstrate that hyper-localised, concerted Spo11 DSBs—double-cuts—also form, separated by just ~33 to over 100 base pairs. Remarkably, double-cut lengths vary with a periodicity of ~10.5 base pairs, conserved in yeast and mouse, invoking a model where the orientation of adjacent Spo11 molecules is fixed relative to the DNA helix—a proposal supported by in vitro DNA-binding properties of the Spo11 core complex. Deep sequencing meiotic progeny identifies recombination scars consistent with repair initiated from gaps generated by adjacent Spo11 DSBs. Collectively, these results revise current thinking about the mechanics of Spo11-DSB formation and expand upon the original concepts of gap repair during meiosis to include DNA gaps generated by Spo11 itself.

Mre11 nuclease-independent Spo11 oligos

During meiotic DNA cleavage, Spo11 becomes covalently attached to each DNA end via a 5′ phosphotyrosyl link[1,2]. Repair of Spo11-capped DSB ends proceeds via a combination of the Mre11 endo- and exonuclease activities[1,5,6] (Fig.1a, *left*). Importantly, the DNA

[6]Corresponding Authors: m.neale@sussex.ac.uk; valerie-corinne.garcia@inserm.fr.

Author contributions

M.J.N. and V.G conceived the project and prepared the manuscript. D.J., V.G, C.C.B., and M.J.N performed physical analysis of Spo11-DCs. M.C. prepared and analysed whole genome recombination maps with B.L. advising upon the mechanistic interpretation. T.C. and M.J.N mapped and analysed Spo11-oligo library data. C.C.B. and S.K. contributed protein biochemistry and provided critical mechanistic interpretations. All authors helped write the manuscript.

Competing interests statement

The authors declare no competing interests

Europe PMC Funders Author Manuscripts

bases in proximity to Spo11 are resistant to degradation[6], thereby generating stable Spo11-oligonucleotide (Spo11-oligo) complexes (Fig.1a). Immunoprecipitation and sequencing of Spo11 oligos has enabled Spo11 activity to be mapped within yeasts[3,7], plants[8] and mammals[4].

When radiolabelled and separated by gel electrophoresis, *S. cerevisiae* Spo11 oligos form two prominent populations[5] of ~8–12 nt and ~25–35 nt (e.g. Fig.1b, c, **blue arrows**), the significance of which has so far remained unclear. Advances in assay sensitivity additionally reveal a faint ladder of slower-migrating species (Fig.1b, **left, red arrows**). Intriguingly, unlike canonical Spo11 oligos, this ladder was retained upon deletion of *SAE2* (Fig.1b, **right, red arrows**), an activator of Mre11 endonuclease activity[9].

High-resolution analysis of deproteinised Spo11 oligos demonstrated that the ladder has ~10 nt periodicity—matching the helical pitch of B-form DNA—from ~33 nt to >100 nt (Fig.1c, Extended Data Fig.1a,b), which partially overlapped with the longer class of canonical Spo11 oligos[5]. The ladder also arose in *rad50S* and Mre11 nuclease-defective strains (Extended Data Fig.1c)—which, like *sae2*Δ, cannot remove Spo11 from DSB ends[5,6]. Importantly, the ladder depended on Spo11's catalytic activity (Extended Data Fig.1d), demonstrating that it is a specific product of Spo11-DSB formation. We estimate that this ladder is ~19% as abundant as the canonical Spo11 oligos detected in wild-type cells (Fig.1d, Extended Data Fig.1e,f).

We considered that this ladder of larger Spo11-oligo complexes might be independent of Mre11 nuclease activity because it arises via Spo11 cleaving DNA at adjacent locations[10,11] (Fig.1a, ***right***). If this idea is correct, we reasoned that the average distance between adjacent DSBs might increase in a strain where catalytically inactive (but DNA-binding competent) Spo11 (Spo11-YF) is expressed alongside wild-type Spo11. Whilst the effect was relatively modest, this prediction was met (Extended Data Fig.1g), suggesting that inactive Spo11 complexes can act as competitive inhibitors of adjacent catalysis.

We henceforth term this ladder as hyperlocalised Spo11 double-cuts (Spo11-DCs). Temporal analysis in wild type revealed Spo11-DCs to arise concomitantly with Spo11 oligos, but to remain visible for longer (Fig.1c), suggesting greater stability or an extended time of formation. Importantly, the very detection of Spo11-DCs suggests they are refractory to Mre11-dependent nucleolytic conversion into Spo11 oligos.

The ATM kinase has an evolutionarily-conserved function in negatively regulating DSB numbers, in part via DSB interference which inhibits coincident DSB formation at adjacent chromosome positions[4,10–15]. While global Spo11-DSB formation increases only ~2-fold upon deletion of *TEL1* (the *S. cerevisiae ATM* orthologue[10,11]), coincident DSB formation in hotspots ~2 kb apart increases up to ~14-fold[10]. Because Spo11-DCs also arise from adjacent DSBs (but presumably formed within the same hotspot and at much closer range), we reasoned that Spo11-DC formation may also be subject to Tel1-dependent inhibition. However, similar to the Spo11 oligos of canonical size, the intensity of the Spo11-DC banding pattern increased only ~2-fold in *tel1*Δ (Fig.1d) with unaltered size distribution and kinetics compared to wild type (Extended Data Fig.1h), suggesting that Tel1 has no

greater role in inhibiting adjacent hyperlocalised Spo11-DC than it does in general DSB suppression.

Spo11-DCs in whole-genome libraries

We investigated whether Spo11-DCs are present within the deep-sequencing Spo11-oligo libraries from wild-type and *tel1*Δ strains[11]. Importantly, size selection during Spo11-oligo library generation[11] excludes short Spo11-oligos and may also underestimate larger Spo11-DCs. Nevertheless, remapping of paired-end sequence reads revealed not just the prominent class previously characterised (20–40 nt), but also peaks in the size distribution at ~43, ~53, ~63, and ~73 nt (Extended Data Fig.2a) consistent with Spo11-DC sizes detected on gels.

Spo11 cut sites display DNA sequence biases spanning ±15 nt around the cleavage site[3] (Extended Data Fig.2b). We reasoned that if molecules ascribed to Spo11-DCs indeed arise from two coincident DSBs they should have this bias at both termini and not just the 5′ end. To test this prediction, we stratified Spo11-oligo length and assessed the average base composition around the 3′ end of each molecule (Extended Data Fig.2c-e). As expected, 27-nt molecules within the major Spo11-oligo peak displayed a 3′ signature unlike Spo11 (Extended Data Fig.2c), attributed to sequence biases of Mre11 nuclease[3]. By contrast, 43-nt Spo11 oligos revealed a clear echo of the Spo11 sequence bias at their 3′ ends (Extended Data Fig.2d), not seen when assessing one of the troughs in the length distribution profile, 39 nt (Extended Data Fig.2e).

Next, sequence reads were filtered to include only overlapping forward-strand (Franklin) and reverse-strand (Rosalind) read pairs with 5′ and 3′ ends offset by 2 nt (the overhang created by Spo11 cleavage[3,16])—with the expectation that this will enrich Spo11-DCs (Fig.2a, top). Remarkably, this exercise enhanced both the Spo11-like sequence bias at the 3′ ends, and the ~10 nt periodicity (Fig.2a-b, Extended Data Fig.2f-h), even though the filter did not include a constraint on read length. A strong peak also arose at 33 nt (Fig.2a), consistent with the shortest Spo11-DC detected physically (Fig.1c), whereas molecules <30 nt were depleted (Extended Data Fig.2g). These results strongly support the conclusion that the periodic Spo11-oligo lengths >30 nt, detected both within sequencing libraries and physically, are indeed Spo11-DCs. Thus, we define Spo11-DCs in these libraries as filtered Spo11 oligos >30 nt (Fig.2a, Extended Data Fig.2i).

We hypothesized that the absence of Spo11-DCs <30 nt in length (Fig.1c) is due to the inability of two DSB-forming complexes to assemble on DNA in such close proximity. To test this idea, we examined recombinant Spo11-Rec102-Rec104-Ski8 complexes, which can bind tightly and noncovalently to dsDNA ends in vitro[17]. Spo11 complexes were incubated with dsDNA fragments of varying length and assessed for their ability to bind one versus both ends (Fig.2c, Extended Data Fig.2j). Remarkably, double-end binding was efficient for DNA molecules that were 30–34 bp or 24–25 bp, but not sizes in between (Fig.2d, Extended Data Fig.2j). We infer that adjacent Spo11 complexes clash sterically at distances below ~30 bp if oriented in the same direction, but this clash is alleviated by a relative rotation of 180° (half a helical turn; Fig.2d). Yet, because Spo11-DCs <30 nt are not detected on gels, and are depleted from filtered Spo11-oligo libraries, we propose that adjacent Spo11 complexes

capable of making double cuts must, in vivo, interact with DNA from the same orientation, thereby generating a ladder of Spo11-DCs with periodicity dictated by the helical pitch of DNA, ~10.5 bp.

Mapping Spo11-DCs within DSB hotspots

To determine the relationship between Spo11-DCs and Spo11-DSB hotspots, we generated and compared genome-wide maps of both. Whilst the proportion of Spo11-DCs within each hotspot varied in both wild type and *tel1*Δ, global Spo11-DC frequency was disproportionately associated with regions of stronger Spo11 activity (Extended Data Fig.3a-e). We visualised Spo11-DCs as frequency-weighted arcs linking the 5′ ends of overlapping Franklin- and Rosalind-strand filtered reads (Fig.2e-f, Extended Data Fig.4a). Like Spo11-DSB cleavage sites, Spo11-DCs were distributed non-uniformly within hotspots, with preferred regions for Spo11-DC formation identifiable as peaks of local enrichment (Fig.2e, Extended Data Fig.4b-d, lower panels). Narrower hotspots displayed a simple pattern with centrally focused Spo11-DCs, whereas wide hotspots contained multiple zones of Spo11-DC enrichment (Fig.2f, Extended Data Fig.4c-d).

Spo11-oligo strand disparity

Every DSB is expected to generate two Spo11 oligos, one from each strand, yet many Spo11-oligo sites display strong disparity between strands[3]. Because Spo11 oligos form two prominent size populations[5], and only longer oligos are mapped[3], such disparity is thought to reflect strand-biased generation of long oligos at certain positions[3,5]. However, the molecular explanation for differential Spo11-oligo size, and for strand-biased formation of long oligos at some cleavage sites, has been unclear.

Critically, we detected an unanticipated alternating pattern of strand-biased Spo11-oligo enrichment across hotspots that was spatially associated with Spo11-DC regions. Specifically, left flanks of Spo11-DC peaks appeared enriched for Franklin-strand hits, whereas right flanks were enriched for Rosalind-strand hits (Fig.2e-f, Extended Data Fig.4b-d). This relationship was retained in *tel1*Δ (Extended Data Fig.4e). We assessed the generality of these observations by aggregating strand-specific Spo11-oligo signals centred on the strongest Spo11-DC peak in each hotspot (Fig.2g). Mapped 5′ ends of bulk Spo11 oligos displayed peaks ±20 bp from Spo11-DC centres and were associated with ~2–fold skews towards Franklin-strand hits on the left and Rosalind-strand hits on the right (Fig.2h).

Moreover, when considering the positions of individual Spo11-DC 5′ ends, a similar strand disparity in local Spo11 oligos was observed, resulting in a ~4-fold differential towards Franklin on the left and Rosalind on the right (Extended Data Fig.5a-g). Importantly, no net disparity was observed when assessing 5′ ends of the entire Spo11-oligo library (Extended Data Fig.5h-j), indicating that these strand-specific patterns are a feature of genomic sites that generate Spo11-DCs. However, the relative scarcity of Spo11-DCs compared to Spo11 oligos means that, whilst Spo11-DCs are associated with cleavage sites that show disparity, they cannot be the cause.

Spo11-DCs arise during mouse meiosis

To investigate whether Spo11-DC formation is evolutionarily conserved we applied the overlapping filter to mouse SPO11 oligos obtained from wild-type and *Atm*$^{-/-}$ spermatocytes[4]. In mouse, the SPO11-oligo size distribution is distinct between wild type and *Atm*$^{-/-}$ [14], with prominent populations at ~10–20 nt and ~25–30 nt in the wild type, and additional larger molecules (25–60 nt) more abundant in *Atm*$^{-/-}$ (Extended Data Fig.6a). Whilst less efficient than in yeast, overlap filtering disproportionately retained putative mouse SPO11-DCs in the 30–60 nt range, but in *Atm*$^{-/-}$ only, where subtle periodic peaks of enrichment at ~32 nt, 43 nt, and 53 nt emerged (Fig.2i, Extended Data Fig.6b), similar to *S. cerevisiae* Spo11-DC sizes. Such filtered SPO11 oligos (>30 nt) were rare in wild type, being at least 10-fold more abundant within annotated hotspots in *Atm*$^{-/-}$ (Extended Data Fig.6c).

In *Atm*$^{-/-}$, bulk SPO11-oligo signals >30 nt in length were offset from hotspot centres in a strand-specific manner (Extended Data Fig.6d), reminiscent of the asymmetric left–right Franklin–Rosalind disparity observed within *S. cerevisiae* hotspots. Moreover, similar to *S. cerevisiae*, but in *Atm*$^{-/-}$ only, plotting SPO11-oligo patterns revealed alternating regions of strand disparity that were associated with the locations of Spo11-DC arcs (Fig.2j, Extended Data Fig.6e-h). Collectively, these observations suggest that Spo11-DC formation is evolutionarily conserved, but may be subject to differential regulation by Tel1 and ATM.

Spo11-DCs lead to gap repair

We reasoned that Spo11-DCs might be repaired as short double-strand gaps, leading to obligate transfer of genetic information from the uncut donor DNA to the gapped molecule (Fig.3a). To test this idea, we analysed meiotic recombination ‘scars’ in *msh2Δ* hybrid yeast containing ~65,000 nucleotide-sequence polymorphisms[18,19]. In *msh2Δ* cells, recombination events initiated by a single DSB are expected to generate heteroduplex DNA (hDNA) retained as segments of 5:3 marker segregation[18]. By contrast, 6:2 segregation should be enriched within events initiated by gaps (Fig.3a).

As previously noted[18], a large fraction (~37%) of recombination events in *msh2Δ* cells contain segments of 6:2 segregation (Fig.3b). Because 6:2 segments can also be generated by mechanisms other than adjacent Spo11 DSBs[19], we isolated events with 6:2 segments 30–150 bp long (~25% of total events), consistent with Spo11-DC sizes determined physically, and then further filtered for events with flanking hDNA in trans orientation—the pattern expected for gap repair (category A; ‘compatible’; Fig.3b-c, Extended Data Fig.7a-g). Importantly, limited polymorphism density (1 per ~200 bp) impedes detection of shorter 6:2 segments (Extended Data Fig.7h), and also causes many events to lack hDNA information (category C; ‘ambiguous’, 12% of total). Nevertheless, ~3% of events were category A, and ~5% were category B (cis hDNA; “incompatible”; Fig.3b-c). Loss of Tel1 had no impact on the relative proportion of categories A–C, but increased the frequency of events with 6:2 segments >150 bp (Extended Data Fig.7a,i), consistent with a loss of DSB interference increasing the probability of coincident DSBs at adjacent hotspots[10].

6:2 segments that do not arise from Spo11-DCs may arise from the DNA-nicking action of Mlh1-Mlh3 and Exo1[19–21]. Consistent with this idea, abrogation of this nicking activity specifically reduced category B and C, whereas category A was unaffected (Fig.3d-e, Extended Data Fig.7a). We further reasoned that 6:2 segments arising from gap repair should correlate with the location of DSB initiation, whereas, for example, nick translation during repair may arise away from the initiation site. Thus, we computed both the frequency of 6:2 segments overlapping Spo11 hotspots (Fig.3f), and the density of Spo11 oligos or Spo11-DCs (Extended Data Fig.7j-k) arising within 6:2 segments, for all categories. By all metrics, category A segments displayed the greatest association with Spo11 activity. These conclusions were unchanged in *tel1*Δ (Extended Data Fig.8l-n). Collectively, these analyses support the view that a subset of recombination events arise from gap repair of Spo11-DCs.

Discussion

Although double-strand gap repair assays were instrumental in establishing the DSB repair model more than 35 years ago[22], direct Spo11-dependent gap formation in vivo was not anticipated, nor does it feature within current models. Indeed, despite an excess of Spo11 protein compared to DSBs[5], and the DNA interaction surface of Spo11 being <30 bp[3,23], it has been generally assumed that only a single DSB is created within any given hotspot even though most meiotic hotspots are many times wider[3,4] (100-1000 bp).

Challenging this idea, we have demonstrated that Spo11 can cleave adjacently at close proximity within hotspots, and provided evidence that such Spo11-DCs yield double-strand gaps that cause obligate transfer of genetic information. Our conclusions broadly agree with observations from the Klein laboratory (pers. comm.), and extend interpretations of recombination patterns in mismatch repair-defective strains[18,19] and Spo11-oligo sizes in *tel1* and *Atm*$^{-/-}$ mutants[4,11].

The ~10-bp periodicity of Spo11-DC sizes is intriguing. Similar periodicity occurs for DNase I cleavage of DNA wrapped around histones[24]. Within nucleosomal DNA, periodicity arises due to only one face of the helix being solvent exposed, with the alternating major-minor groove pattern repeating once per helical turn. Importantly, however, DSB hotspots have open chromatin structure, and stable nucleosomes occlude Spo11 access to DNA in vivo[3], similar to other topoisomerase family members[25]. Therefore, we consider that something other than nucleosomes restricts Spo11-DC endpoints to these periodic positions.

We envision a mechanism wherein multiple Spo11 proteins assemble with other pro-DSB factors creating a platform that enables concerted Spo11-DSB formation (Fig.4). In this model, periodic spacing arises by restricting Spo11's access to the same face of the DNA helix, with preferred cleavage opportunities found only every helical turn (~10.5 bp) due to stiffness of B-form DNA over these short distances. This model explains why the minimum Spo11-DC size is ~33 bp despite our DNA-binding observations indicating that adjacent Spo11 complexes can come as close as 24 bp in vitro when their relative orientation is not constrained. Moreover, a multimeric DSB-forming 'machine' made up of many catalytic centres may explain the apparent excess of Spo11 protein[5], and why *SPO11/spo11-Y135F*

catalytic heterozygosity has little effect on overall DSB formation despite an expectation of negative dominance[23]. Finally, the periodicity strongly suggests that the two DSBs that generate a Spo11-DC happen concertedly as part of the same DSB-forming reaction.

In our model, the alternating strand disparity in Spo11-oligo libraries arises from the Spo11 platform protecting a subset of DSB ends from the nuclease activities of Mre11[6]. For DSB ends within the platform, or facing inwards from the platform edges, this protection may yield Spo11-DCs as well as the canonical Spo11 oligos of larger sizes (Fig.4a-b, Extended Data Fig.8a). By contrast, outermost DSB ends are less protected, yielding Spo11 oligos vulnerable to more extensive digestion by Mre11, in turn leading to their under-representation in Spo11-oligo maps that omit short oligos. Thus, we propose that the two prominent size classes of Spo11 oligo detected physically[5] arise from differential sensitivity to nucleolytic degradation caused by asymmetric interactions of hotspots with the chromosome axis (Fig.4a-b, Extended Data Fig.8a).

In broad terms, our model for a surface-bound Spo11 platform is compatible with how Spo11 and its essential cofactors interact on the chromosomal axis in vivo[26–31], and with DNA-dependent assembly of the Spo11 core complex with its cofactors in vitro[17,32]. Whilst we draw the axis in a planar form, our ideas do not exclude a model where the platform and hotspot DNA curve or writhe in concert with one another, and it is upon the exposed surface of such a structure that DSB formation occurs. However, the fact that we detect no major DNA sequence skews towards more flexible A or T bases in the centre of Spo11-DC fragments (Extended Data Fig.2h) disfavours the idea that DNA is subject to significant localised bending forces during Spo11-DC formation. Nevertheless, Spo11-DC sizes are 1–2 bp larger than expected for relaxed B-form DNA, suggesting that modest underwinding of the DNA helix may be integral to Spo11-DSB formation.

Tel1 inhibits coincident DSB formation at adjacent hotspots[10,12]. Yet, surprisingly, *tel1Δ* had a relatively modest impact on Spo11-DC formation despite each Spo11-DC also being formed by a pair of adjacent DSBs. We propose that DSBs can arise concertedly within a DSB-active region (creating Spo11-DCs) before the inhibitory effect of Tel1 can act—promoted by the high-density Spo11 array (Extended Data Fig.8b). By contrast, our exploration of mouse *Atm*$^{-/-}$ Spo11-oligo libraries suggests that whilst Spo11-DC formation within hotspots may be evolutionarily conserved—consistent with molecular analysis of recombination in *Atm*$^{-/-}$ spermatocytes[33]—the regulation by Tel1/ATM is not. Specifically, whilst Spo11-DCs were abundant in wild-type yeast, SPO11-DCs in wild-type mouse were infrequent and increased much more upon *Atm* deletion than upon deletion of *TEL1* in yeast (~10-fold, Extended Data Fig.6c vs ~2-fold, Extended Data Fig.3b). These differences suggest that SPO11 double cutting in mouse is subject to more efficient direct inhibition by ATM. Alternatively, because ATM controls resection (via Mre11), this difference may arise due to increased sensitivity of Spo11-DCs to Mre11 in mouse.

Our observations deepen understanding of the meiotic recombination pathway to include concerted Spo11-DSB formation, therein providing a glimpse into how the elusive biochemistry of Spo11 works in vivo.

Methods

Strains and culture methods

S. cerevisiae strains used in this study are listed in Extended Data Table 1. Synchronous meiotic cultures were grown using standard methods. Briefly, YPD cultures (1% yeast extract, 2% peptone, 2% glucose) were diluted 100-fold into YPA (1% yeast extract, 2% peptone, 1% K-acetate) and grown vigorously for 14 h at 30°C. Cells were collected by centrifugation, washed once in water, resuspended in an equal volume of prewarmed 2% K-acetate containing diluted amino acid supplements, and shaken vigorously at 30°C. For mapping meiotic recombination patterns in hybrid octads, SK1 and S288c haploid parents were mated for 8–14 hours on YPD plates as described[34]. Cells were washed and incubated in sporulation media at 30°C with shaking, and tetrads were dissected after 72 hours. To generate octads, dissected spores were allowed to grow for 4-8 hours on YPD plates until they had completed the first post-meiotic division, after which mother and daughter cells were separated by microdissection and allowed to grow for a further 48 hours. Spore clones were subsequently grown for 16 hours in liquid YPD prior to genomic DNA isolation using standard techniques[10]. Only octads generating eight viable progeny were used for genotyping by deep sequencing.

Spo11-oligo and Spo11-DC physical analysis

Spo11-oligo complexes were isolated by immunoprecipitation using anti-FLAG antibody from 10 ml aliquots of sporulating cells harvested at the indicated timepoint(s), and labelled with alpha-32P cordycepin triphosphate using Terminal deoxynucleotidyl transferase (Fermentas) as described[35], and separated by 7.5% SDS-PAGE, or treated with Proteinase K (Fisher) for 1 hour at 37°C prior to overnight precipitation in 90% ethanol at -80°C, then denatured in 1x formamide loading dye and separated on 19% denaturing urea/PAGE in 1x TBE. Where indicated, deproteinised samples were further treated with 300 nM recombinant mammalian TDP2 for 30 minutes at 30°C[35] prior to electrophoresis, to remove residual phosphotyrosyl linked peptides, thereby enabling an accurate estimate of Spo11-oligo DNA length. Radioactive signals were collected on phosphor screens, scanned with a Fuji FLA5100 and quantified using ImageGauge software. Uncropped gel images are included in the Source Data file. For analysis of Spo11-oligo species by Western, SDS-PAGE gels were transferred to PVDF membrane, blocked with 5% non-fat dry milk (NFDM) / 1x TBST, and incubated with anti-FLAG-HRP at 1:1000 in 1% NFDM / 1xTBST.

Quantification of Spo11-DCs (signals >30 nt) relative to Spo11-oligo signals (<31 nt) was performed using ImageGauge software v4.0 (Fuji) after normalising loading between samples based on Spo11 protein abundance as estimated using non-proteolysed samples separated by SDS-PAGE and detection via anti-FLAG Western blotting. Based on prior[5], and our current observations, only a very small fraction of Spo11 participates in Spo11-DSB formation (we estimate <5%). Thus, any remaining material that does not enter the gel in *sae2*Δ will only represent a very tiny fraction, and therefore we believe will only have a very minor effect on the normalisation we have performed. When quantifying Spo11-oligo and Spo11-DC signals on denaturing acrylamide gels, the inherent overlap between the size range of Spo11-DC and Spo11-oligo signals, and varied contribution of lane background,

means that our stated values are only estimates within a range of possible values. Due to different choices that can be made when setting background subtraction for the Spo11-DC molecules, we calculate both minimum and maximum estimates and present the mid-value of these. Extended Data Fig.1e-f provides examples of our quantification methodology. Importantly, whilst we consider that this is an unbiased manner to estimate Spo11-DC abundance, we believe that setting the background at its minimum leads to an overestimate of the Spo11-DCs, and therefore results in modest inflation of the reported mid value. Quantifications are also unable to assess the relative impacts (and potential differential rates) of turnover of Spo11-DCs relative to Spo11-oligos, nor the potential conversion of Spo11-DCs into canonical Spo11-oligos via the natural resection process catalysed by Mre11–Sae2. However, the facts that Spo11-DCs are detected at all, and the relatively similar abundance of Spo11-DCs in both resection proficient (wild type) and resection deficient (*sae2*Δ) *S. cerevisiae* cells, suggests that Spo11-DC may be relatively inert and certainly refractory to the rapid pathway of resection that arises at regular Spo11-DSB ends.

To estimate the fraction of events containing a Spo11-DC, we apply the simplifying assumption that, at most, only a single Spo11-DC arises per event such that each Spo11-DC makes four Spo11-oligos (two internal, two external) whereas each single Spo11 DSB makes only two oligos (two external). Applying these assumptions, the estimated fraction of events containing a Spo11-DC simplifies to the equation: Spo11 oligos>30 nt / Spo11 oligos<30 nt.

Assuming there is, at most, only one Spo11-DC per recombination initiation event, and taking into account uncertainties in quantification, we estimate that up to ~19% of events contain a Spo11-DC. The lower frequency estimated for gap repair in recombination data (~3%) may reflect overestimates in Spo11-DC quantifications, and uncertainties in ascribing recombination events to gap repair, inability to detect intersister recombination, and effects of limited polymorphism density (1 per ~200 bp), which reduces by more than half the chance of detecting gaps <150 bp (Extended Data Fig.7h).

Remapping of *S. cerevisiae* Spo11-oligo libraries

S. cerevisiae Spo11-oligo libraries[11] were aligned to the Cer3H4L2 reference genome using Bowtie2, with identical GLOBAL and LOCAL mapping parameters: -X1000 --no-discordant --very-sensitive --mp 5,1 ----np 0. Cer3H4L2 is identical to the sacCer3 reference genome (R64-1-1), with the addition of two ectopic insertions: 1173 bp of hisG sequence inserted at the *LEU2* locus at position 91965, and 3037 bp of LEU2 sequence including 77 bp of associated unidentified bacterial sequence[36] at position 65684. Before mapping, reads were trimmed to remove adapters and trailing G or C bases introduced during library preparation using Perl (OligoTrim.pl). Specifically, 5′ ends of Read1 were trimmed using Perl to remove the following sequences at the first 9 or 8 bases of each read: NNNNNCCCC (or NNNNNCCC if prior sequence not found). Read1 3′ ends were trimmed to truncate before any GGGGAGAT (or GGGAGAT if prior sequence not found) sequences should they be present. Similarly, Read2 5′ ends were trimmed to remove leading CCCC (or CCC if prior sequence not found) sequences, and 3′ ends truncated before GGGGNNNNNAGAT (or GGGNNNNNAGAT if prior sequence not found) sequences. In each case, the need to

trim the NNNNN string arises from the use of custom barcoded adapters during library preparation[11]. The AGAT string is the reverse complement of the first 4 bp of the universal Illumina adapter. Following trimming and paired end alignment, the Read1 5′ base is expected to have a high probability of being a true Spo11-oligo 5′ end, and the terminal mapped base a true 3′ end. Nevertheless, some ambiguity is impossible to avoid due to inherent uncertainties in the number of terminal rG bases added during library preparation[11]. Resulting SAM files were processed via terminalMapper (https://github.com/Neale-Lab) using the 'DOUBLE' setting, generating 1 bp resolution histogram files of 5′ Spo11 oligo ends mapping to either Franklin or Rosalind strands of the genome. Additional "CoordinateAB" files report frequencies, strand, and position of molecules with unique 5′ and 3′ ends, enabling filtering for overlapping pairs. In these files, the 3′ reported is 2 nt more distal than the actual 3′ end such that it corresponds to any putative 5′ end on the complementary strand. As such, and because the AB coordinates listed include the first and last mapped base, the raw oligo lengths are 1 nt shorter than obtained by subtracting the B coordinate from A. Upper panels plot raw strand-specific data normalised to hits per million mapped reads (HpM). In lower panels, strand-specific data was smoothed using a 51 bp sliding Hann window.

Remapping of mouse Spo11-oligo libraries

Mouse *Atm* $^{-/-}$ SPO11-oligo libraries[4], were trimmed in a similar way. Specifically, 5′ ends of Read1 were trimmed using Perl to remove the following sequences at the first 9 or 8 bases of each read: NNNNNCCCC (or NNNNNCCC if prior sequence not found). Read1 3′ ends were trimmed to truncate before any GGGG (or GGG if prior sequence not found) sequences should they be present. Similarly, Read2 5′ ends were trimmed to remove leading CCCC (or CCC if prior sequence not found) sequences, and 3′ ends truncated before GGGG (or GGG if prior sequence not found) sequences. Resulting SAM files were processed with terminalMapper as above, but additionally processed to expand the called 5′ and 3′ ends by ±1 bp in a manner that averages the frequency of reads mapped at a specific pair of coordinates equally across the nine resulting combinations. This additional step was introduced due to our lower confidence in the accuracy of 5′ and 3′ mapped coordinates based on our inability to trim many reads based on their expected sequence composition. This process increases the likelihood of isolating overlapping pairs of SPO11 oligos (which otherwise requires a precise 2 bp offset between 5′ and 3′ ends of the overlapping pair), at the expense of 9-fold lowered absolute frequency of any given molecular coordinates.

Filtering of Spo11-oligo libraries to isolate Spo11-DC compatible molecules

To enrich for putative Spo11-DC molecules in both *S. cerevisiae* and mouse libraries, Spo11-oligos were discarded unless an overlapping partner, mapped with 2 bp offset at both ends, was present within the library. Spo11-DCs were then conservatively reported as twice the minimum frequency of either the Franklin or Rosalind oligo (whichever was lower). Additionally, for quantitative analyses and the plotting of arc-diagrams, all oligos <31 nt were discarded, justified by our physical gel-based analysis of Spo11-DCs in *sae2*Δ cells, and the periodic enrichment of Spo11-DCs for oligos >30 nt. The estimated proportion of Spo11-DCs within the Spo11-oligo libraries (obtained via our overlapping filter, and thresholded for molecules >30 nt) is lower, but still in rough agreement (~4.6%), to that

estimated on gels in both wild type and *sae2*Δ cells (8-25%). The discrepancy in these proportions likely arises from inaccuracies in quantifying the weak signals on gels (we estimate that these Spo11-DC molecules represent less than ~1% of the total Spo11 protein), limitations of the filtering process, and the size selection undertaken during Spo11-oligo library preparation.

Calculating DNA sequence composition at 5′ DSB ends

DNA sequence orientated 5′→3′ around 5′ cleavage sites for the relevant library fraction was aggregated using seqBias (Perl, v5.22.1; https://github.com/Neale-Lab) and plotted as fractional base composition at each base for the top strand.

Calculating strand disparities

When Spo11 5′ ends are mapped to a reference genome, the 2 bp overhang at cleavage sites translates to a 1 bp offset for the mapped 5′ base on each strand. Thus, to compute the relative strand disparity (ratio of the frequency of mapping, in hits per million, HpM, on each strand at a given cleavage site), the mapped Rosalind coordinates were first shifted by -1 bp. Before calculating disparity ratios, 0.01 HpM were added to both denominator and numerator to avoid errors arising when one of the values was zero. Net left–right disparity for individual Spo11-DC molecules was defined as: $(\text{Franklin}_{\text{left}} / \text{Rosalind}_{\text{left}}) / (\text{Franklin}_{\text{right}} / \text{Rosalind}_{\text{right}})$.

Analysis of 6:2 recombination patterns in *msh2*Δ hybrid octad data

All analysed meioses included *msh2*Δ, which abrogates mismatch repair, thereby enabling the retention of heteroduplex DNA (hDNA) strand information, which then segregates (becoming homoduplex) in the first post-meiotic cell division (octad stage). Whole-genome DNA libraries for each haploid member of every octad were independently prepared and barcoded using Illumina NexteraXT according to manufacturer's instructions, and sequenced at ~16-plex on a MiSeq using 2 x 300 bp paired-end reads, obtaining a minimum of at least 20x average genome coverage. Paired-end mapping, genotyping of the ~65,000 SNP and Indels present in each of the eight haplotypes in each octad, and HR event calling across each octad were performed as described[34] using publicly available scripts (https://github.com/Neale-Lab/OctadRecombinationMapping), generating Event Tables listing position, type, and detailed information describing each isolated HR event (see Source Data File). An integral stage of event calling is the partition of the octad into genomic segments of identical adjacent marker segregation (e.g. 5:3, 6:2, 4:4, etc). As with prior work[19], an inter-event merging threshold of 1.5 kb was used—that is, a minimum of at least 1.5 kb of 4:4 marker segregation was necessary between two adjacent regions of non-4:4 marker segregation for the region to be recorded as two independent HR events.

In HR events from *msh2*Δ mutants, segments of 6:2 marker segregation are not expected from simple models of DSB repair, but may arise due to either secondary nicking by an associated nuclease[19], or from initiation by an adjacent pair of Spo11-DSBs[18]—referred to here as a Spo11-DC. In order to focus on events that may have arisen from putative gap repair—and compatible with the range of Spo11-DC sizes detected physically (Fig.1, Extended Data Fig.2)—events were filtered to include only those containing a

6:2 segment, and then further categorised depending upon the estimated length of the 6:2 segment, whether there was more than one 6:2 segment within the event (single vs multi), and the marker segregation patterns flanking the 6:2 segment (categories A, B, C; Fig.3, Extended Data Fig.7). Category A events are compatible with gap repair because flanking heteroduplex DNA patterns are in trans orientation, whereas category B events are incompatible because the flanking heteroduplex DNA patterns are in cis orientation. Category C events lack the flanking heteroduplex DNA patterns necessary to assign the event.

Because the perceived length (and visibility) of any segment is affected by local variant density, categorisations are inherently uncertain, but represent our reasonable estimates. Firstly, 6:2 segments were excluded from consideration when the maximum possible length of a 6:2 segment was <30 bp, and when the minimum length of a 6:2 segment was >150 bp. Whilst we don't exclude the possibility that concerted Spo11-DSBs separated by more than 150 bp might arise (see Fig.1c and Extended Data Fig.1a for examples of larger periodic species resolved on gels), we favour the view that the more separated DSBs are, the more likely they will behave as two independent DSBs[10]—each with two recombination-active DNA ends—and thus less likely to generate products compatible with a simple model of gap repair. Additionally, most events were classified as Category C (ambiguous) due to lack of useful flanking heteroduplex information. Nevertheless, when possible, flanking heteroduplex segregation patterns were used to exclude some events (Category B, incompatible)—for example, when the flanking hDNA was in *cis*, rather than *trans* orientation (Fig.3c). In other instances, strand polarity information—inferred from the overall phasing of the haplotypes based on NCO *trans* hDNA patterns present elsewhere in the octad[19]—were used to aid classification. NCOs without phasing information were classified as Category C (ambiguous). Despite these uncertainties, the fact that the fraction of events falling into Category A is unaffected by deletion of *EXO1*, *MLH1* or *MLH3* (which when mutated abrogate the generation of 6:2 events arising from nicking[19], see below), whereas Category B and C are reduced ~3-fold and ~7-fold, respectively, by these mutations, provides confidence in the validity of our classifications. Precise categorisation rules are presented below:

NCO events: NCO 6:2 segments flanked by hDNA tracts in *trans* orientation (e.g. CN4; Extended Data Fig.7g) were considered highly compatible because the hDNA segments are suggestive of repair synthesis tracts. NCO 6:2 segments flanked by hDNA in *cis* orientation were classed as incompatible. NCO 6:2 segments with hDNA on only one side were classified as compatible if the hDNA pattern matches the known phasing of the strands (e.g. CN3; Extended Data Fig.7f). In this latter case, we infer that the missing information on one side of the event may be absent due to lack of variant coverage. One-sided NCO events with incorrect strand phasing were classified as incompatible. One-sided NCO events where the strand orientation could not be phased were classed as ambiguous.

CO events: There are two places a 6:2 segment can appear; in the centre of the CO exchange (e.g. CC1-CC4; Extended Data Fig.7b-c) or offset from the CO exchange (e.g. CC5, CC7; Extended Data Fig.7d-e), but involving one of the two chromatids already involved in the CO, and falling within the 1.5 kb inter-event threshold. Central CO gaps were considered

compatible if there were *trans* hDNA patterns either side of the 6:2 segment. However, unlike the case with NCOs, the *trans* patterns can be across two chromatids if the 6:2 segment is at the CO point (e.g. CC1; Extended Data Fig.7b). In this latter case, the strand orientation phasing was taken into account, and for all COs where the phasing of both chromatids was known, the hDNA patterns were confirmed to be in *trans* (e.g. CC5; Extended Data Fig.7d). CO 6:2 segments where strand orientation phasing of either chromatid was not known but yet display a *trans* hDNA-like pattern were retained in the compatible category (e.g. CC1-2; Extended Data Fig.7b). COs containing offset 6:2 segments were classed as compatible when the gap had full flanking *trans* hDNA (e.g. CC5; Extended Data Fig.7d), or half-hDNA that was in the correct phased orientation (e.g. CC7; Extended Data Fig.7e). COs containing offset 6:2 segments with hDNA in the incorrect orientation (where phasing was possible) were classified as incompatible. Note that the analysis included all CO events, including complex ones showing bi-directional conversions that involve at least two initiating DNA lesions on two non-sister chromatids (e.g. CC3, CC5, CC7; Extended Data Fig.7c-e)[19]. Such events were included when the outcomes of the initiating lesions could be reasonably anticipated.

Overlap of 6:2 segments with annotated hotspots

To calculate hotspot overlap, 6:2 segments up to but not including the non-6:2 flanking markers, were tested for their intersection with the coordinates of a list of previously annotated Spo11-DSB hotspots[11], generating a binary, yes/no result for each 6:2 segment within each category (A–C). Proportions were then calculated and reported. For these analyses the Spo11-oligo datasets utilised were obtained from SK1 nonhybrid diploids[11].

Measuring Spo11 activity within, and surrounding 6:2 segments

To assess the correlation between 6:2 segment locations and local, population average, Spo11-DSB activity, each HR event was partitioned into 6:2 and non-6:2 segments, and the observed amount of Spo11-oligo signal[11] falling within each partition calculated. For this analysis, 6:2 segments were defined as the region up to but not including the non-6:2 flanking markers. Expected Spo11 signal for the 6:2 segment was calculated based on the fraction of total Spo11-oligo signal expected to fall within this segment were Spo11-DSBs arising uniformly across the entire event region. Finally, observed/expected ratios were calculated for each 6:2 segment within each category (A–C), and plotted as individual points on a log2 scale. Box-and whisker plots indicate median (horizontal bar), upper and lower quartiles of the range (box), and minimum and maximum points within 1.5-fold of interquartile range (whiskers). For these analyses the Spo11-oligo datasets utilised were obtained from SK1 nonhybrid diploids[11]. Similar analysis was also performed using the observed amount of Spo11-DC detected in each 6:2 segment (Extended Data Fig.7k), leading to broadly similar conclusions. However, the absence of Spo11-DCs in many of the category B and C 6:2 segments prevented their analysis, artefactually inflating the global ratio reported for these categories, and thereby decreasing the difference when compared to Category A segments (which more frequently overlap with both Spo11 oligos and Spo11-DCs—and are thus not artefactually inflated).

Analysis of *exo1*Δ, *mlh1*Δ and *mlh3*Δ data in *msh2*Δ background

To determine the impact of abrogating the putative nicking activity promoted by Exo1, Mlh1 and Mlh3, previously published datasets[19] were analysed using the methods described above. Due to the limited number of repeats (two meioses for each genotype), and the expected similarity in phenotype of each null mutation[19], these data were pooled and analysed in aggregate ("*exo-mlh*Δ") in order to increase statistical power. Whilst there may be specific differences in phenotype that arise from each individual mutation (for example, *exo1*Δ specifically alters resection tract length, and *mlh1*Δ results in longer hDNA tracts[19]), the data were aggregated due to the overall similarity of phenotypes described[19] and their genetically inferred role at the same step in the class I crossover resolution pathway.

Method for handling incomplete octads

Occasionally we encountered octads where a mother-daughter pair had not been separated correctly, resulting in a 'septad'—identified by a chromatid displaying no visible hDNA information. In these cases, we removed the affected chromatids from our analysis, and thus any HR events arising on these chromatids were subtracted from the total event count.

Electrophoretic mobility shift assays

DNA substrates were generated by annealing complementary oligos (sequences below) generating 2-nt 5′ TA overhangs at both ends. Oligos were mixed in equimolar concentrations (10 mM) in STE (100 mM NaCl, 10 mM Tris-HCl pH 8, 1 mM EDTA), heated and slowly cooled. Substrates were 5′ end labeled with gamma-32P-ATP and T4 polynucleotide kinase and purified by native polyacrylamide gel electrophoresis. The core complex containing Spo11, Rec102, Rec104 and Ski8 from *S. cerevisiae* was purified from baculovirus-infected insect cells[17].

Binding reactions (20 μl) were carried out in 25 mM Tris-HCl pH 7.5, 7.5% glycerol, 100 mM NaCl, 2 mM DTT, 5 mM MgCl2 and 1 mg/ml BSA with 0.5 nM DNA and the indicated concentration of core complexes. Complexes were assembled for 30 minutes at 30 °C and separated on a 5% Tris-acetate-polyacrylamide/bis (80:1) gel containing 0.5 mM MgCl2 at 200 V for two hours. Gels were dried, exposed to autoradiography plates and revealed by phosphorimaging. Uncropped gel images are included in the Source Data file. Notably, the Spo11 core complex binds preferentially to 2-nt 5′ overhangs (<0.5 nM for 50% binding) compared to blunt ends (~3 nM for 50% binding), and most weakly to hairpins (>10 nM)[17]. Thus, we interpret the supershift observed at just ~3 nM to indicate robust coincident binding to both 2-nt 5′ DNA overhangs.

Oligonucleotides for substrate preparation

TAGCAATGTAATCGTCTATGACGTGTCATAGCGC	34 bp - Top
TAGCGCTATGACACGTCATAGACGATTACATTGC	34 bp - Bottom
TAGCAATGTAATCGTCTATGACGTGCATAGCGC	33 bp - Top
TAGCGCTATGCACGTCATAGACGATTACATTGC	33 bp - Bottom
TAGCAATGTAATCGTCTATGACGTGATAGCGC	32 bp - Top

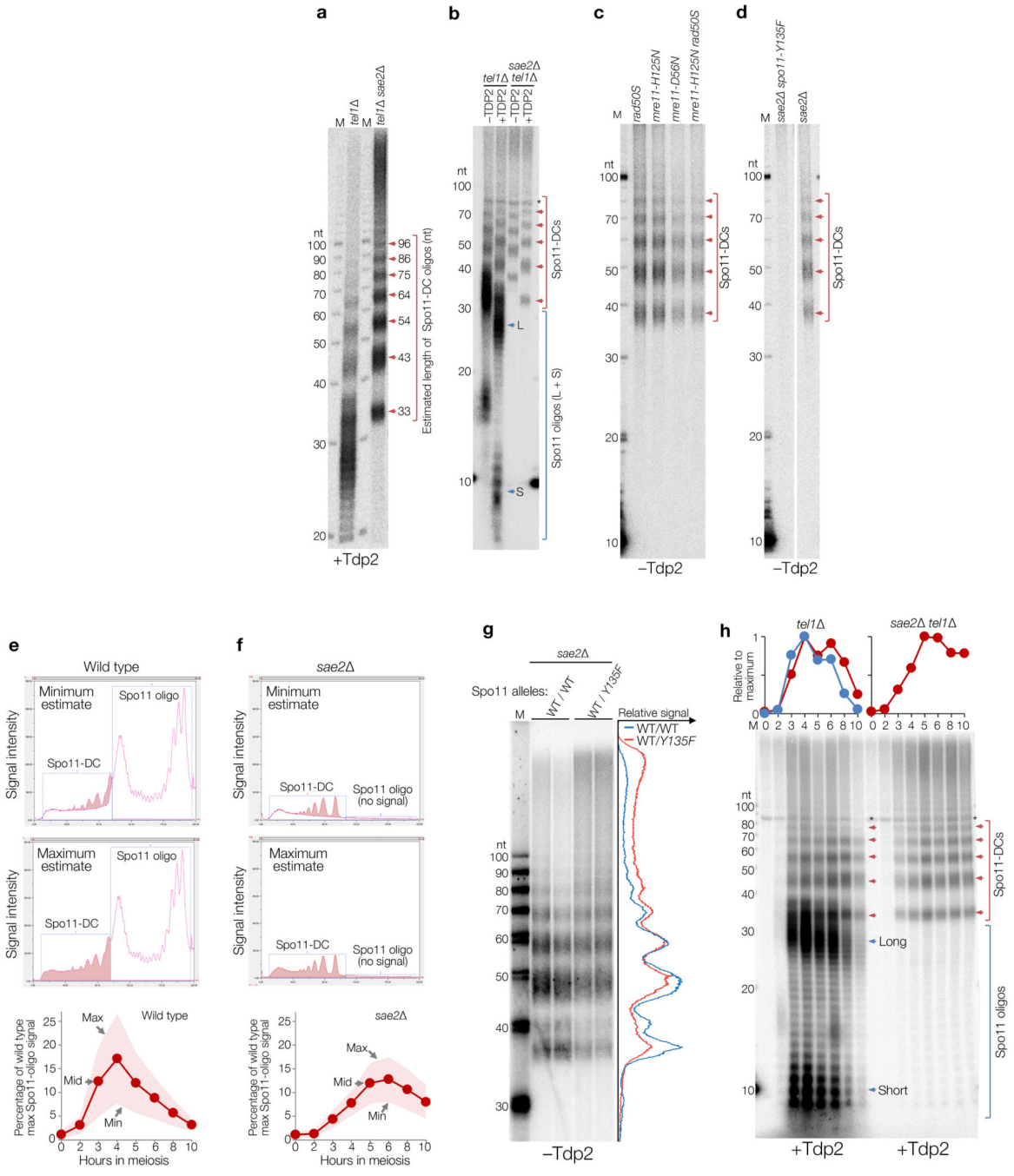
TAGCGCTATCACGTCATAGACGATTACATTGC	32 bp - Bottom
TAGCAATGTAATCGTCTATGACGTGTAGCGC	31 bp - Top
TAGCGCTACACGTCATAGACGATTACATTGC	31 bp - Bottom
TAGCAATGTAATCGTCTATGACGTTAGCGC	30 bp - Top
TAGCGCTAACGTCATAGACGATTACATTGC	30 bp - Bottom
TAGCAATGTAATCGTCTATGACGTTAGCG	29 bp - Top
TACGCTAACGTCATAGACGATTACATTGC	29 bp - Bottom
TAGCAATGTAATCGTCTATGACGTTAGC	28 bp - Top
TAGCTAACGTCATAGACGATTACATTGC	28 bp - Bottom
TAGCAATGTAATCGTCTATGACGTTAG	27 bp - Top
TACTAACGTCATAGACGATTACATTGC	27 bp - Bottom
TAGCAATGTAATCGTCTATGACGTTA	26 bp - Top
TATAACGTCATAGACGATTACATTGC	26 bp - Bottom
TAGCAATGTAATCGTCTATGACGTT	25 bp - Top
TAAACGTCATAGACGATTACATTGC	25 bp - Bottom
TAGCAATGTAATCGTCTATGACGT	24 bp - Top
TAACGTCATAGACGATTACATTGC	24 bp - Bottom
TAGCAATGTAATCGTCTATGACG	23 bp - Top
TACGTCATAGACGATTACATTGC	23 bp - Bottom
TAGCAATGTAATCGTCTATGAC	22 bp - Top
TAGTCATAGACGATTACATTGC	22 bp - Bottom

Statistics and Reproducibility

Biologically independent experimental repeats of each genotype analysed for Spo11-oligo and Spo11-DC formation via polyacrylamide gel were undertaken at least three times with similar results. Relevant quantifications, with the number of biological repeats (n) that were averaged together, are presented in Fig 1d. Representative cropped gel images are presented in Figures 1b-c, 2c, and Extended Data Figures 1a-d, 1g-h, 2j, with corresponding uncropped images provided in the Source Data File. Quantifications presented in Figures 1c, 2d, and Extended Data Figures 1g-h, refer to the individual representative gel that is presented. *S. cerevisiae* Spo11-oligo data analysis uses equally mixed pools of five wild type and five *tel1*Δ samples spanning timepoints 4, 5, and 6 hours in meiotic prophase. Analysis of *S. cerevisiae* meiotic recombination patterns via F1 hybrid SK1 x S288c octad analysis pools data from nine *msh2*Δ octads, ten *tel1*Δ *msh2*Δ octads, and two octads each of *mlh1*Δ *msh2*Δ, *mlh3*Δ *msh2*Δ, and *exo1*Δ *msh2*Δ datasets[19].

Europe PMC Funders Author Manuscripts

Extended Data

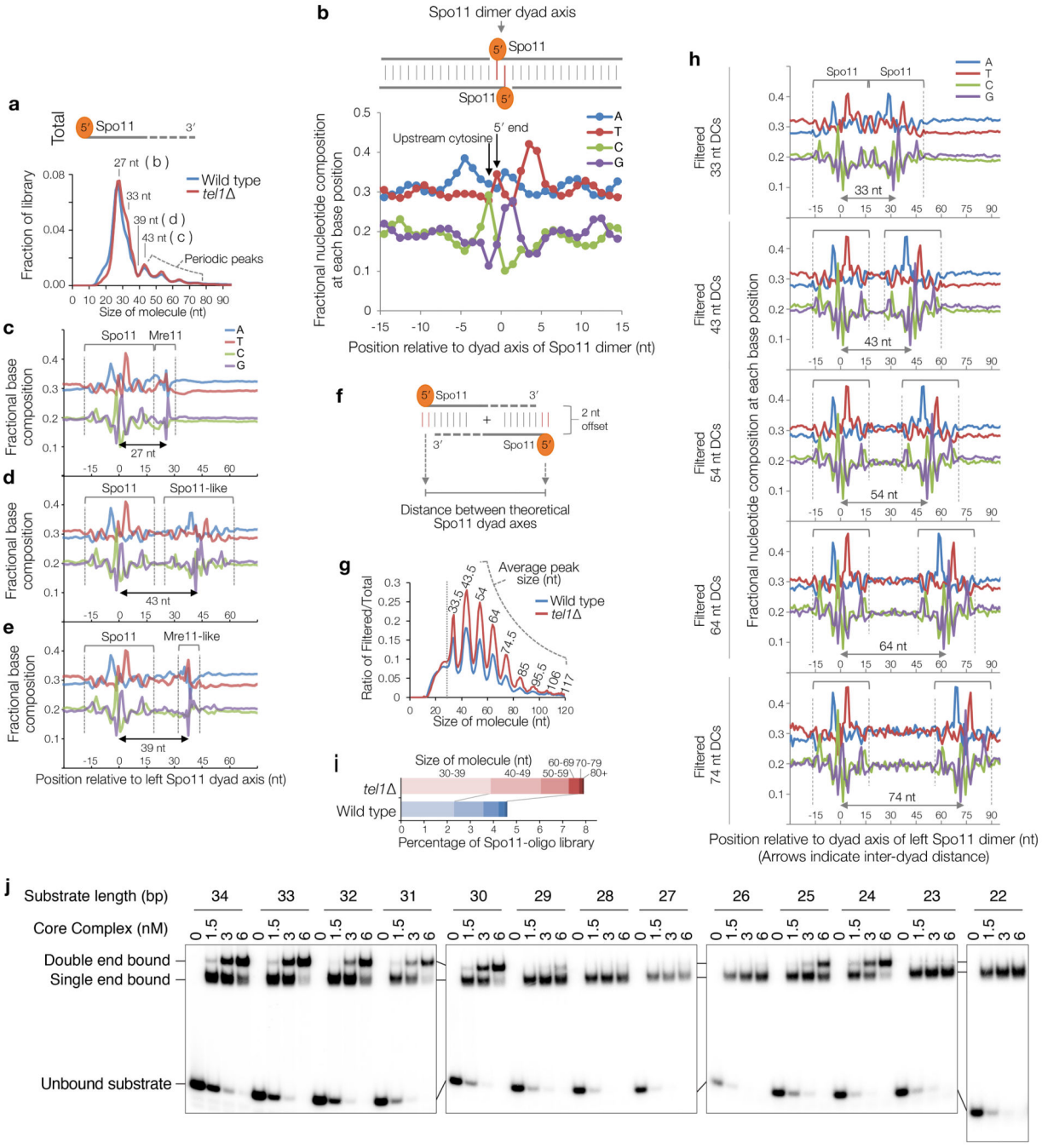


Extended Data Figure 1. Spo11-DC sizing, quantification and genetic analysis.

a-h, Immunoprecipitated *S. cerevisiae* Spo11 oligos and Spo11-DCs isolated from meiotic extracts of the indicated mutants (at 5 hours, or indicated number of hours, after induction of meiosis) were radiolabelled with chain-terminating 3′-dATP using terminal transferase and separated on 19% denaturing PAGE following digestion with proteinase K. Where indicated, samples were also treated with mammalian Tyrosyl DNA phosphodiesterase 2, TDP2[37],

which removes the residual Spo11 peptide that is left after proteinase K treatment[35], thereby permitting accurate estimation of the Spo11-DC oligo length. In the absence of TDP2 digestion, the residual 5′-linked Spo11 peptide retards migration of Spo11 oligos and Spo11-DCs by the equivalent of ~6-8 nt (**b-d, g**). A 10-nt ladder (also radiolabelled with 3′-dATP) is included in each gel to permit accurate sizing. Spo11-DCs arose independently of single or dual mutations in the MRX complex (*rad50S*, *mre11-H125N*, *mre11-D56N*) that abrogate endonuclease activity[1,5,6,38–41] (**c**), and were abolished by homozygous mutation of the Spo11 active site (*spo11-Y135F*) (**d**). Loading normalisation was not performed on samples in panel **c**, therefore differences in Spo11-DC abundance do not convey information. In **c-d** Spo11-DCs were not treated with TDP2 leading to slower migration. In (**e-f**), representative 4-hour lane traces of sequencing gels shown in Fig.1c are shown using two modes of background subtraction (upper, and centre panels), with resulting maximum, minimum, and mid Spo11-DC quantifications (lower panels). Shaded areas in upper panels are the area being quantified. Shaded area in lower quantification data show the range between maximum and minimum, as indicated in the figure. Quantified average mid values are reported in Fig.1d (min–max range of 8–25%). Further quantification details are provided in Methods. *SPO11*/*spo11-Y135F* heterozygous diploids (**g**) display an altered Spo11-DC oligo size distribution (biological duplicate lanes of each are presented alongside averaged intensity trace). Analysis of Spo11-oligo and Spo11-DC intermediates at hourly timepoints during meiotic prophase in the absence of Tel1 (**h**).

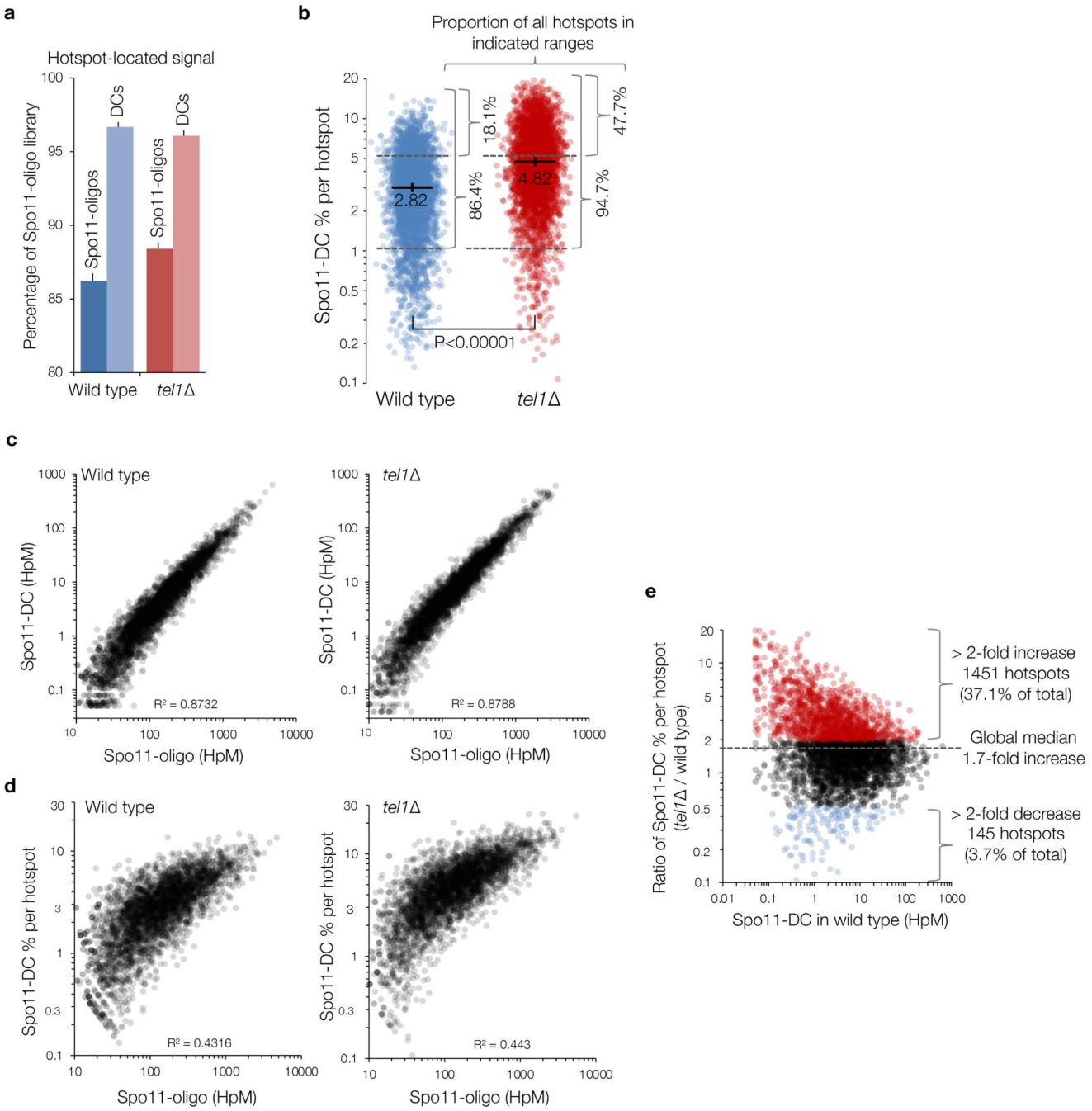
Europe PMC Funders Author Manuscripts



Extended Data Figure 2. Biased sequence composition around Spo11-DC 5′ ends.

Spo11-oligos[11] isolated from the indicated *S. cerevisiae* strains were remapped using paired-end Bowtie2 alignment. **a,** Size distribution of total Spo11 oligos in wild type and *tel1*Δ strains. Periodic peaks in the distribution are indicated, including a subtle shoulder at 33 nt, consistent with the Spo11-DC sizes detected on gels. **b,** Cartoon depicting Spo11 dimer staggered cuts (top) and mean nucleotide composition surrounding the Spo11 cleavage site (bottom) (see Methods). Population-averaged features of Spo11 cleavage sites include preferred cleavage 3′ to a C nucleotide and flanking A/T skews[3,42]. **c-e,** Nucleotide

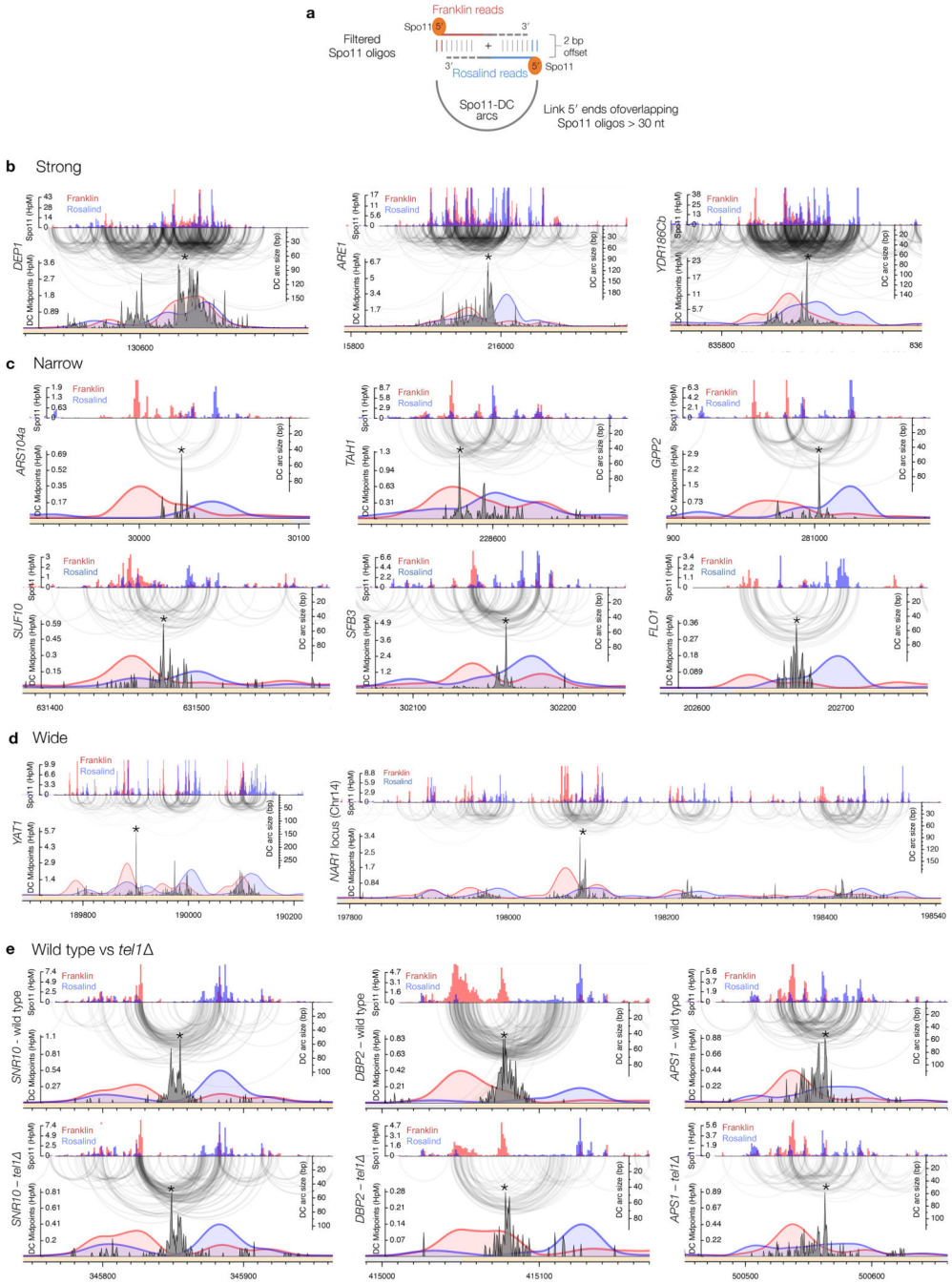
composition of Spo11 oligos of the indicated size was computed for each base position, revealing Spo11 signature at both ends (**d**), or Spo11 signature at 5′ end plus Mre11 signature at 3′ end (**c, e**). **f-i,** Spo11-DC were filtered out from total Spo11-oligo libraries based on overlapping molecules sharing 5′ and 3′ coordinates with a precise 2 nt offset (**f**). The theoretical dyad axis of each Spo11 dimer (at each end of the molecule) is indicated. Due to the rotational symmetry of cleavage, the distance between such dyad axes is identical to the filtered oligo length. In (**g**), the ratio of filtered to total Spo11 oligos is plotted as a function of molecule length. Because molecules <30 nt were not detected on Spo11-oligo gels in a *sae2Δ*, we infer that retention of some molecules <30 nt is due to fortuitous artefactual overlap of canonical Spo11-oligos (~27 nt). Such filtered molecules were therefore excluded from all bioinformatic analyses. In (**h**), the mean nucleotide composition of filtered Spo11 oligos of the indicated size (the sizes presented are peaks in the filtered size distribution) was computed for each base position and plotted relative to the inferred dyad axis of cleavage of the leftmost Spo11 DSB, revealing signature nucleotide skews characteristic of Spo11 at both the 5′ and 3′ ends. No base skews were observed in the central regions of each molecule, arguing against a major influencer of Spo11-DC formation being localised DNA bending, which is expected to be favoured by an AT-rich base composition. In (**i**), the percentage of total Spo11-oligo library in the filtered (Spo11-DC) fraction is plotted for the indicated size ranges. As defined, Spo11-DCs make up ~4.6% and 7.9% of the total Spo11-oligo pool in wild-type and *tel1*Δ strains, respectively, consistent with the *tel1*Δ-dependent increase measured by our physical analysis. In absolute terms, these values are presumably lower than our gel-based estimates due to size selection during library preparation, the stringency of the filtering, and inaccuracies in quantifying Spo11-DCs on gels, which we estimate make up less than 1% of the total cellular Spo11 protein (see Methods). **j,** *In vitro* DNA mobility shift assay as described[17]. Spo11 core complex (Spo11, Rec102, Rec104, Ski8) was incubated with double-strand DNA substrate of different lengths with 2 nt TA overhang on both ends. Based on prior experiments[17], the robust supershift observed at ~3 nM is interpreted to indicate double-end binding by the Spo11 core complex. Quantification is provided in Fig.2d. Whilst the in vitro assay involves heterotetrameric Spo11 complexes (i.e. a Spo11 core-complex monomer), we assume that similar binding characteristics will take place in vivo involving octameric complexes (i.e. Spo11 core-complex dimers).



Extended Data Figure 3. Spo11-DC composition of *S. cerevisiae* DSB hotspots in wild type and *tel1*Δ.

a, Percentage of total Spo11 oligos and filtered Spo11-DCs that arise within annotated DSB hotspots. Overall, nearly all (~95%) Spo11-DCs map within hotspots, more so than total Spo11 oligos (~86%), suggesting that Spo11-DCs are more prevalent where Spo11 activity is strongest. **b,** Percentage of total Spo11 oligos that are Spo11-DCs, plotted for every hotspot (n=3910). Although the proportion of Spo11-DCs within each hotspot varied widely from <0.1% to >10% of the Spo11-oligo signal, the majority (~86%) of hotspots

displayed a Spo11-DC proportion of at least 1%, and about one fifth (~18%) of hotspots displayed Spo11-DC proportions over 5%. In *tel1*Δ, the fraction of hotspots falling into these categories increased to ~94%, and ~47%, respectively, consistent with the median fraction of Spo11-DCs per hotspot being ~1.7-fold greater. P values indicate two-tailed Kruskal-Wallis H-test. **c-d,** Quantitative correlation between filtered Spo11-DC frequency (**c**), or percentage of Spo11-DCs within each hotspot (**d**), and total Spo11-oligo frequency for all DSB hotspots, in wild type (left) and *tel1*Δ (right). Whilst Spo11-DC frequency correlated positively with total Spo11-oligo counts, the relationship was nonlinear, such that Spo11-DCs were observed disproportionately more frequently within the strongest hotspots. **e,** Comparison between *tel1*Δ and wild type of the percentage of total Spo11-oligo signal within each hotspot that is classified as a Spo11-DC. These ratios are stratified on the X-axis by the Spo11-DC frequency in wild type cells, and ratios are coloured to indicate those hotspots where the proportion of Spo11-DCs is at least 2-fold increased (red) or 2-fold decreased (blue) in *tel1*Δ relative to wild type. Note that although Spo11-DCs were globally more frequent in *tel1*Δ compared to wild type, this relationship was not uniform across all hotspots.



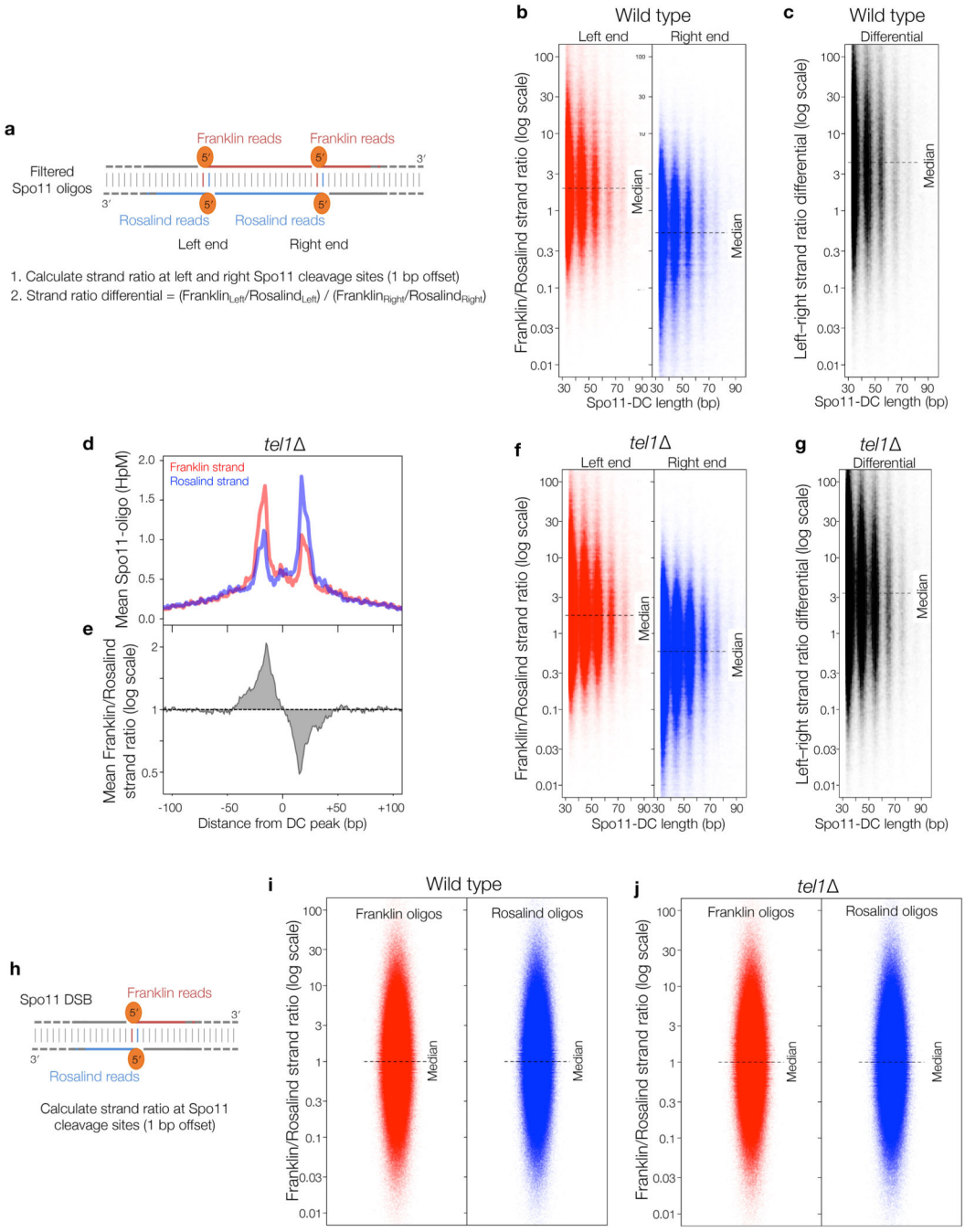
Extended Data Figure 4. Fine-scale patterns of *S. cerevisiae* Spo11-DCs within DSB hotspots.

a, Spo11-DC arcs link the 5′ ends of overlapping Franklin- and Rosalind-strand filtered reads. For all bioinformatic analyses, only overlapping read pairs >30 nt are considered because this is the minimum length of Spo11-DCs detected physically (Fig.1c). We believe that enrichment of some shorter overlapping pairs arises from the artefactual overlap of canonical oligos (<30 nt) within dense hotspot regions. **b-e,** Arc diagram depiction of Spo11-DCs mapped across example hotspots encompassing strong (**b**), narrow (**c**), and wide (**d**) classes, presented as in Fig.2e-f. Top panel: unfiltered strand-specific Spo11 oligos

Europe PMC Funders Author Manuscripts

(Franklin-strand, red; Rosalind-strand, blue; HpM, hits per million mapped reads). Arcs (grey-scale frequency-weighted) link 5′ ends of each Spo11-DC. Lower panel: smoothed unfiltered strand-specific Spo11-oligos, overlaid with frequency histogram of Spo11-DC midpoints (grey). The left flanks of Spo11-DC peaks are enriched for Franklin-strand hits, whereas the right flanks are enriched for Rosalind-strand hits. This relationship was visualised most easily at narrow, low frequency hotspots where Spo11-DC patterns were less complex. In (**e**), wild-type and *tel1*Δ data are compared for the same hotspots. Whilst *TEL1* deletion does have subtle effects on the pattern and abundance of both Spo11 oligos and Spo11-DCs, it did not alter the asymmetric pattern of Spo11-oligo strand disparity that is associated with regions of preferential Spo11-DC formation. Please note that in all plotted Spo11-oligo and Spo11-DC maps, Rosalind-strand signals are shifted by 1 bp to the left so that differences between the abundance of F- and R-mapping reads at individual cleavage sites can be more directly compared.

Europe PMC Funders Author Manuscripts

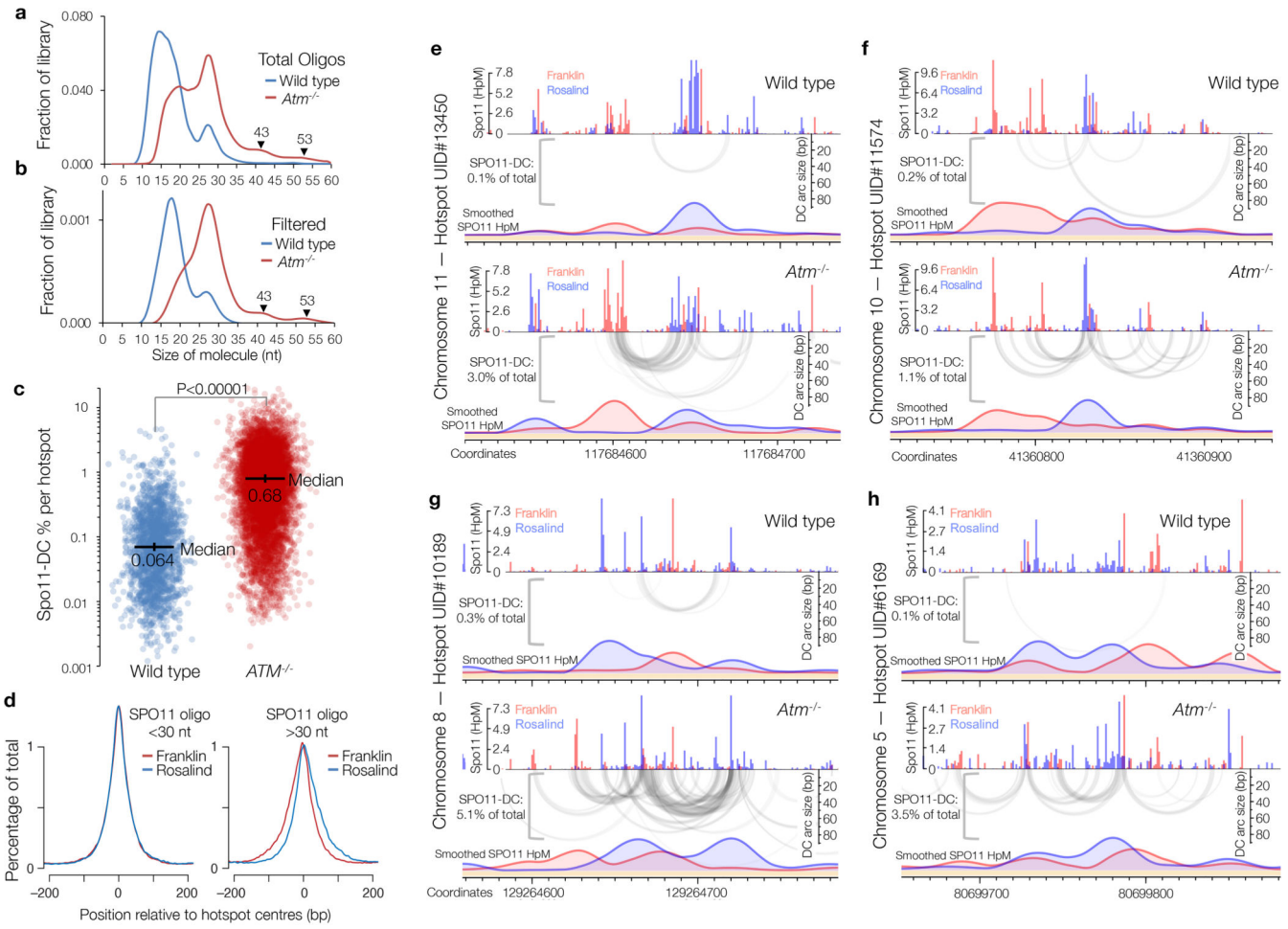


Extended Data Figure 5. Global analysis of strand disparities at Spo11-DC termini in *S. cerevisiae* wild type and *tel1Δ* cells.

a, Explanatory cartoon for the calculation of Spo11-DC strand ratio and strand-ratio differential (Left/Right). See Methods for further details. **b-g,** Strand ratios of Spo11 oligos at Spo11-DC sites in wild type (**b-c**) and *tel1*Δ (**d-g**) cells. Average strand-specific Spo11-oligo signal (**b, f**), and strand ratio (**c, g**), centred upon the strongest Spo11-DC midpoint within every DSB hotspot (n=3910). Strand ratio (Franklin/Rosalind total Spo11-oligo HpM) was computed at the left and right 5′ end of every unique Spo11-DC molecule

(**a**), stratified by length. Strand-ratio differential (**c, g**) indicates the fold difference in the ratios when comparing the left and right 5′ ends of each Spo11-DC molecule. The relationships described above were unchanged upon *TEL1* deletion, suggesting that the Spo11-oligo patterns are an intrinsic feature of sites where Spo11-DCs are generated, and are not subject to regulation by Tel1. **h,** Explanatory cartoon for strand-ratio calculation for all Spo11 oligos. **i-j,** Strand ratio (Franklin/Rosalind total Spo11-oligo HpM) was computed at the 5′ end of all observed (unfiltered) Franklin- or Rosalind-strand Spo11 oligos (**h**). Unlike at Spo11-DC sites (**b, f**), bulk Spo11 oligos, across all sites, display no net strand disparity in either wild type (**i**) or *tel1*Δ (**j**) strains. HpM = Hits per million mapped reads. Please note that in all plotted Spo11-oligo and Spo11-DC maps, Rosalind-strand signals are shifted by 1 bp to the left so that differences between the abundance of Franklin- and Rosalind-mapping reads at individual cleavage sites can be more directly compared. When considering all Spo11-oligo sites (**h-j**), some degree of strand disparity is a feature of most Spo11-oligo sites (it is a continuum of skew in both directions—some sites are skewed towards Franklin, some towards Rosalind, and some have little or no skew), but, when considered in aggregate, bulk Spo11-oligo sites have no net skew towards Franklin or Rosalind regardless of which strand the Spo11-oligo is considered. By contrast, sites where we detect Spo11-DC formation (**a-g**; which is only a subset of all the sites onto which Spo11 oligos are mapped) display an asymmetric average skew within the total Spo11-oligo Franklin and Rosalind reads, with the left end of Spo11-DCs being sites where the total Spo11-oligo pool is skewed towards Franklin and the right end where the total Spo11-oligo pool is skewed towards Rosalind. Importantly, this analysis utilises the total Spo11-oligo pool, not just Spo11-DC molecules. Thus, the pattern of skews is a global feature of the entire Spo11-oligo pool at sites that form Spo11-DC, and is not a feature that is observed at all Spo11-oligo sites. We interpret these observations to mean that sites where Spo11-DCs form are different, being disproportionately sites of biased strand disparity. Due to their low abundance at any particular site, removing Spo11-DCs from the total pool of Franklin and Rosalind reads has no impact on the strand disparity observed. Finally, consistent with these interpretations, we have found that the degree of disparity at any given site is not predictive of Spo11-DC abundance at that site, further indicating that Spo11-DC formation is not the cause of the disparity. Instead, locations of strand disparity and Spo11-DC are correlated in position, likely because they are influenced by similar properties of DSB hotspots (i.e. a proposed Spo11 platform; see below). Thus, overall, we conclude that asymmetric strand disparity is not unique to specific hotspots, nor to specific mutants, nor caused by Spo11-DCs, but is, instead, an intrinsic feature of the meiotic recombination process that informs the mechanism of Spo11-DSB formation.

Europe PMC Funders Author Manuscripts

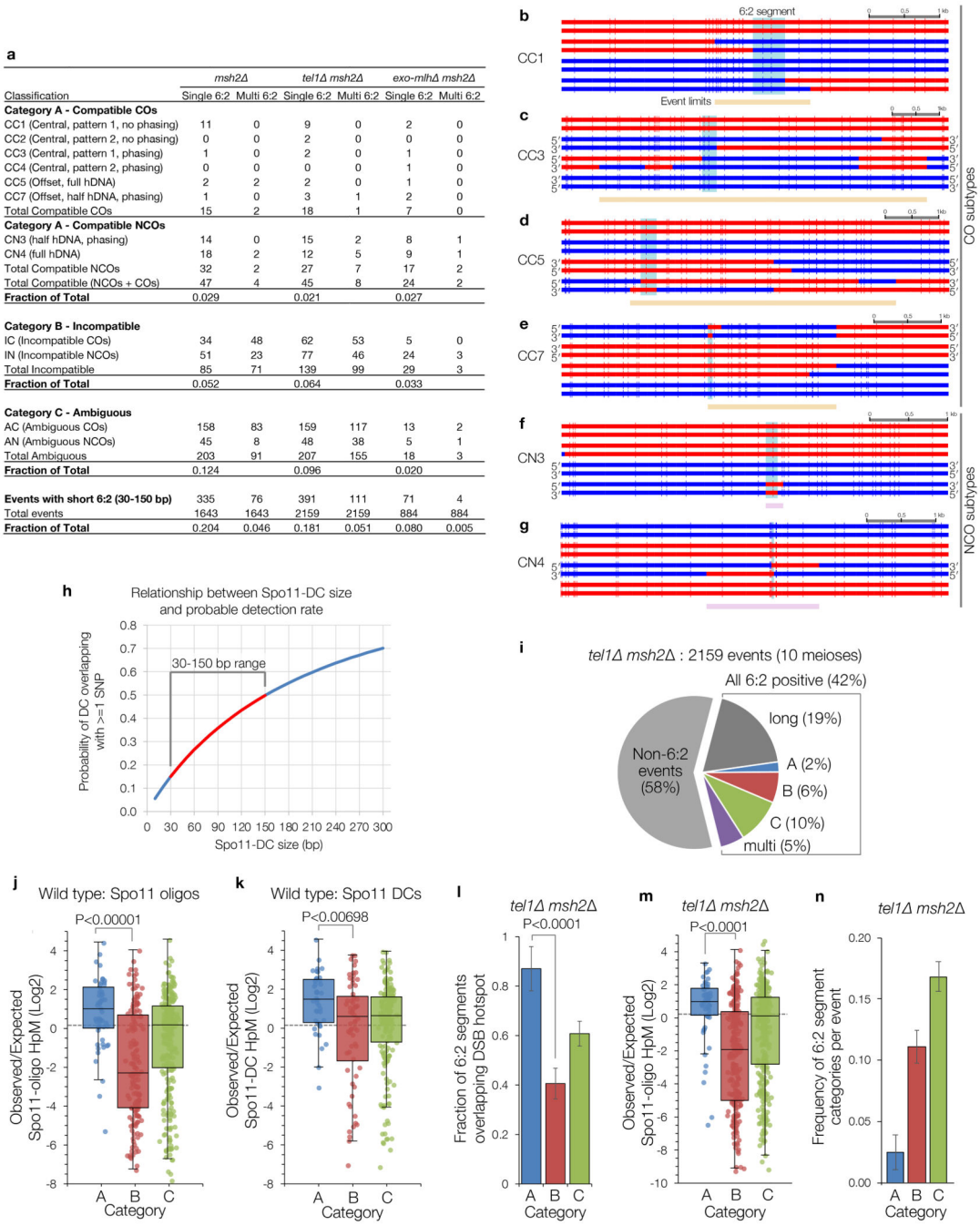


Extended Data Figure 6. Fine-scale analysis of SPO11-DCs within mouse DSB hotspots.

Mouse SPO11-oligo libraries[4] were remapped using paired-end Bowtie2 alignment. **a-b,** SPO11-oligo length distribution of the entire library (**a**), or after applying the 2 bp overlap filter (**b**) as in Fig.2a, upper cartoon. Filtering was less efficient than in yeast (see Methods), retaining numerous molecules <30 nt, that, based upon our analysis in yeast, are likely to be a filtering artefact. Therefore, as for *S. cerevisiae*, SPO11-DCs are defined as filtered molecules >30 nt in length. Filtering retains weak peaks in *Atm* $^{-/-}$ that display a ~10 bp periodicity. **c,** Percentage of total SPO11 oligos that are SPO11-DCs based on overlap filtering, plotted for every hotspot where filtered Spo11-DCs were detected (n=1831 in wild type; n=8010 in *Atm* $^{-/-}$). P value indicates two-tailed Kruskal-Wallis H-test. **d,** Total *Atm* $^{-/-}$ SPO11 oligos were filtered into two size classes then aggregated around ~21,000 hotspot centres revealing a strand-specific disparity for Spo11-oligos >30 nt. **e-h,** Representative arc diagrams of SPO11-DCs (grey-scale frequency-weighted arcs) in wild-type and *Atm* $^{-/-}$ relative to total strand-specific SPO11 oligos (upper, raw; lower, smoothed; red, Franklin strand; blue, Rosalind strand). Percentage of total SPO11 oligos that are SPO11-DCs is indicated. Unlike in *Atm* $^{-/-}$, SPO11-DCs in wild type often did not coincide with strong SPO11-oligo signals, suggesting that some may arise from additional artefacts of the filtering.

a

Classification	*msh2Δ* Single 6:2	*msh2Δ* Multi 6:2	*tel1Δ msh2Δ* Single 6:2	*tel1Δ msh2Δ* Multi 6:2	*exo-mlhΔ msh2Δ* Single 6:2	*exo-mlhΔ msh2Δ* Multi 6:2
Category A - Compatible COs						
CC1 (Central, pattern 1, no phasing)	11	0	9	0	2	0
CC2 (Central, pattern 2, no phasing)	0	0	2	0	0	0
CC3 (Central, pattern 1, phasing)	1	0	2	0	1	0
CC4 (Central, pattern 2, phasing)	0	0	0	0	1	0
CC5 (Offset, full hDNA)	2	2	2	0	1	0
CC7 (Offset, half hDNA, phasing)	1	0	3	1	2	0
Total Compatible COs	15	2	18	1	7	0
Category A - Compatible NCOs						
CN3 (half hDNA, phasing)	14	0	15	2	8	1
CN4 (full hDNA)	18	2	12	5	9	1
Total Compatible NCOs	32	2	27	7	17	2
Total Compatible (NCOs + COs)	47	4	45	8	24	2
Fraction of Total	0.029		0.021		0.027	
Category B - Incompatible						
IC (Incompatible COs)	34	48	62	53	5	0
IN (Incompatible NCOs)	51	23	77	46	24	3
Total Incompatible	85	71	139	99	29	3
Fraction of Total	0.052		0.064		0.033	
Category C - Ambiguous						
AC (Ambiguous COs)	158	83	159	117	13	2
AN (Ambiguous NCOs)	45	8	48	38	5	1
Total Ambiguous	203	91	207	155	18	3
Fraction of Total	0.124		0.096		0.020	
Events with short 6:2 (30-150 bp)	335	76	391	111	71	4
Total events	1643	1643	2159	2159	884	884
Fraction of Total	0.204	0.046	0.181	0.051	0.080	0.005

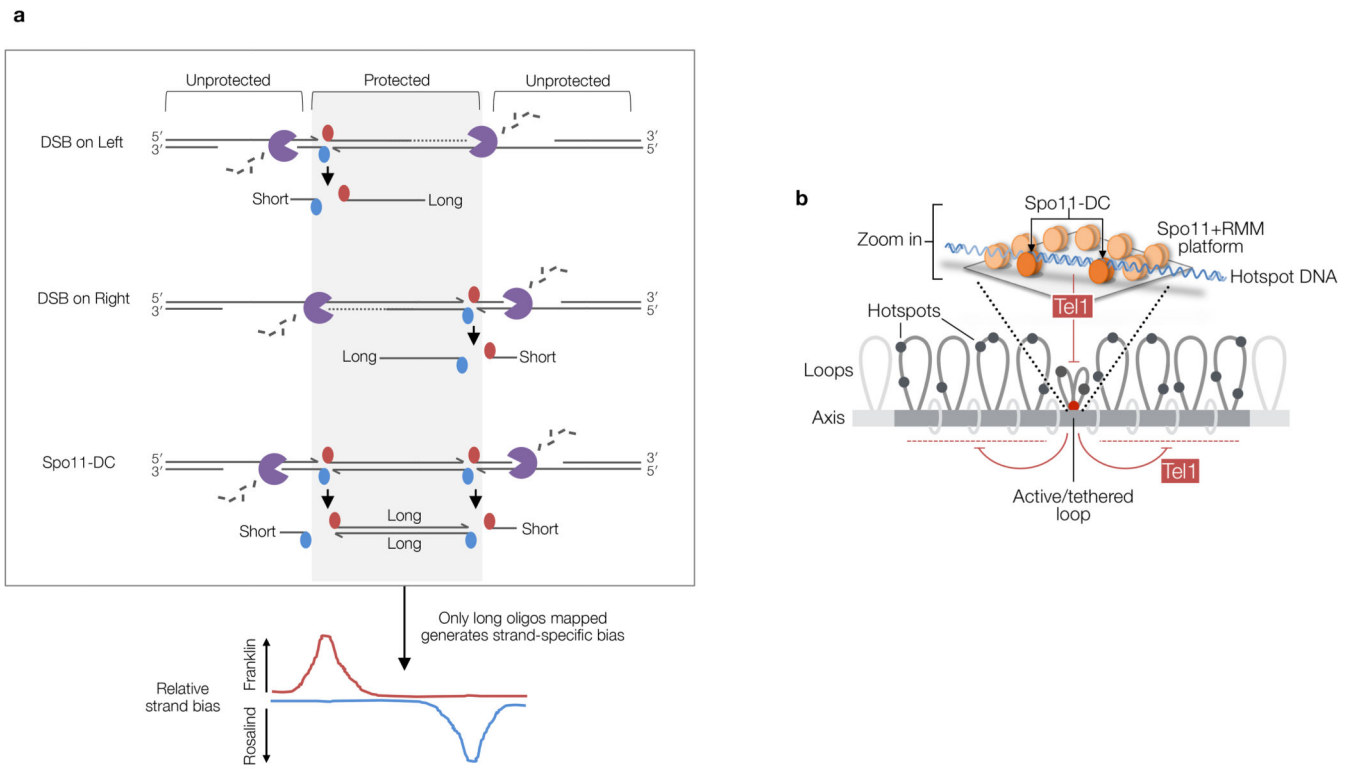


Extended Data Figure 7. Categorisation of *S. cerevisiae* meiotic recombination events containing short 6:2 segments.

a, Summary of frequencies of each subclassification event type for the indicated strains. Only events containing 6:2 segments 30 to 150 bp in length were considered. Category A events are compatible with gap repair because flanking heteroduplex DNA patterns are in trans orientation, whereas category B events are incompatible because the flanking heteroduplex DNA patterns are in cis orientation. Category C events lack the flanking heteroduplex DNA patterns necessary to assign the event. Events were separated into

those with a single or multiple such 6:2 segments. Fractions of total events for each subtype were not calculated for multi events because they frequently contain more than one sub-type. **b-g,** Example event sub-classifications. Genotype calls were made at each marker (vertical line). Adjacent segments of the same genotype are joined with horizontal bars (red or blue) to aid visualisation of patterns. Each horizontal bar is a sequenced haplotype from one meiotic octad. 6:2 segments are indicated in pale blue. Event limits are indicated by beige (crossover) or pink (noncrossover) bars. Orientation of 5′ and 3′ strands are indicated in instances where it was possible to obtain phasing information from noncrossover *trans* events within the event, or from events elsewhere in the octad (see Methods for further details). In (**c**), a second segment of 6:2 segregation was not considered because the minimum length is >1.2 kb. **h**, Relationship between Spo11-DC size and the probability of it overlapping at least one SNP. To estimate probable detection rates of theoretical Spo11-DC of varying size, sliding windows of increasing size were moved across the reference genome, and the number of genetic markers within each window was recorded for each position. As examples, on average, Spo11-DCs 30 bp and 150 bp in size will be detected only 15% and 50% of the time, respectively. We note that due to the non-uniform distribution of both Spo11 DSBs and genetic markers—in particular the slightly greater density of polymorphisms within intergenic regions where Spo11 DSBs most often arise—the probability of detection may be slightly greater than that estimated from the genome-wide polymorphism density. **i**, Quantification of recombination event types in *tel1*Δ *msh2*Δ based on categories presented in Fig.3b-c. Relative proportions of category A–C were unchanged when compared to wild type, but a greater number of events containing 6:2 segments >150 bp were observed. **j-k**, Log2 ratio of observed Spo11-oligo (**j**) or Spo11-DC (**k**) density within each 6:2 segment divided by the Spo11-oligo or Spo11-DC density within the entire event, in *msh2*Δ, for each category. Whilst these analyses broadly agree with each other, the absence of Spo11-DCs in many of the category B and C segments prevented their analysis, artefactually inflating both their global ratio and decreasing the difference when compared to Category A (which overlaps with both Spo11 oligos and Spo11-DCs). **l-n,** *TEL1* deletion has no impact on the association between recombination patterns containing 6:2 segments and Spo11-DSB activity. Fraction of 6:2 segments overlapping hotspots in *tel1*Δ *msh2*Δ (**l**), as for Fig.3f. Log2 ratio of observed Spo11-oligo (**m**) density within each 6:2 segment divided by the Spo11-oligo density within the entire event, in *tel1*Δ *msh2*Δ, for each category, as for (**j**). Unlike deletion of *MLH1*, *MLH3* and *EXO1* (Fig.3e), deletion of *TEL1* had no impact on the relative frequencies of 6:2 segments within categories A–C (**n**). In (**j-k, m**), P values indicate two-tailed Kruskal-Wallis H-test. Individual data points are shown overlaid with box indicating median plus first and third quartiles, and whiskers indicating 1.5x the interquartile range. In (**l, n**), whiskers indicate 95% confidence intervals. In (**l**), P value indicates two-tailed Z-test of proportions. In (**i, l-m**), 2159 events were analysed across 10 biologically independent meiotic samples. In (**j-k**), 1643 events were analysed across 9 biologically independent meiotic samples.

Europe PMC Funders Author Manuscripts



Extended Data Figure 8. Cartoon to explain mapped strand disparity and Tel1-insensitive Spo11-DC formation.

a, Model to account for observed strand bias. Mre11-dependent 3′→5′ exonuclease activity is shown relative to the covalent attachment of Spo11 to 5′ DNA ends. We hypothesise that Spo11 oligos and Spo11-DCs formed within the axis-associated Spo11 platform (grey area) are protected from Mre11 nuclease activity and are therefore efficiently mapped (long Spo11 oligos), while efficient resection in the flanking regions leads to shorter Spo11 oligos that are not mapped. We assume that because the position and frequency of such hotspot–axis interactions will vary from one hotspot to another—and from one cell to another—this will contribute to the substantial variations in position and abundance of both Spo11 oligos and Spo11-DCs within hotspot regions. Whilst it is formally possible that the disparity can alternatively arise from less (rather than more) efficient resection flanking the platform area, leading to unprocessed DSB ends, and thereby a paucity of Spo11-oligo reads in these locations, we consider this unlikely because it would require there to be a high frequency of persistent unresected DSBs, something which is not observed in wild-type cells. **b,** In *S. cerevisiae*, Tel1 has no greater impact on Spo11-DC formation than on global DSB formation (Fig.1d), yet efficiently inhibits DSB formation between adjacent hotspots[10]. Therefore, we propose that DSBs arise concertedly within a DSB-active hotspot region (upper cartoon)—creating Spo11-DCs—before Tel1 can act to inhibit their formation. Such hotspot activation likely arises via the proposed tethering of nucleosome-depleted hotspot DNA to the pro-DSB axis components[27,43–45] (lower panel). Such interactions will enable the formation of both single Spo11-DSBs and/or Spo11-DCs at the tethered locus, either of which will cause Tel1 activation. Once activated, Tel1 inhibits DSB formation at hotspots within the rest of the tethered loop[10], and at any axis-associated DSB hotspots within

adjacent loop regions. Such inhibition may arise via direct inhibition and/or destabilisation of Spo11 and/or other pro-DSB axis components such as Rec114-Mer2-Mei4 (RMM), and/or inhibition of hotspot-axis interactions[10,11,15].

Extended Data Table 1
Strains used in this study.

Unless indicated, all strains are of the SK1 genetic background[46]. For recombination mapping in octads, haploids were mated directly prior to sporulation as described[19,34]. The specific *lys2* mutation is not known, but has been historically present in SK1 background strains. The *ura3* allele contains an insertion of a Ty transposon. When not directly indicated, the *arg4* mutant allele is either *arg4-nsp* or *arg4-bgl*. When not directly indicated, the *leu2* mutant allele is either *leu2*::*hisG* or *leu2Δ* allele. We have no indication that allelic differences at these loci have any influence on the key observations reported in this study. *lys2*, *ura3*, *arg4*, and *leu2* alleles all confer lack of growth on drop-out media lacking the relevant component.

Strain name	Genotype	Origin
VG296	*MATa/alpha, ho::LYS2/", lys2/", ura3/", arg4/", leu2::hisG/", his4X::LEU2,/" nuc1::LEU2/", SPO11-His6-FLAG3-loxP-KanMX-loxP/"*	This study
VG303	*MATa/alpha, ho::LYS2/", lys2/", ura3/", arg4/", leu2/", his4X::LEU2,/" nuc1::LEU2/", SPO11-His6-FLAG3-loxP-KanMX-loxP/", sae2Δ::KanMX6/"*	This study
VG300	*MATa/alpha, lys2/", ura3/", arg4/", leu2/", his4X::LEU2,/" nuc1::LEU2/", SPO11-His6-FLAG3-loxP-KanMX-loxP/", tel1Δ::HphMX4/"*	This study
VG302	*MATa/alpha, lys2/", ura3/", arg4/", leu2/", his4X::LEU2,/" nuc1::LEU2/", SPO11-His6-FLAG3-loxP-KanMX-loxP/", tel1Δ::HphMX4/", sae2Δ::KanMX6/"*	This study
SKY 3935	*MATα; ho::LYS2; lys2; ura3; leu2::hisG; his3::hisG; SPO11-5ProA-his5+sp*	Mohibullah et al, 2017
SKY 3934	*MATα; ho::LYS2; lys2; ura3; leu2::hisG; his3::hisG; SPO11-5ProA-his5+sp*	Mohibullah et al, 2017
SKY 3950	*MATα; ho::LYS2; lys2; ura3; leu2::hisG; his3::hisG; SPO11-5ProA-his5+sp; tel1Δ::kanMX*	Mohibullah et al, 2017
SKY 3951	*MATα; ho::LYS2; lys2; ura3; leu2::hisG; his3::hisG; SPO11-5ProA-his5+sp; tel1Δ::kanMX*	Mohibullah et al, 2017
MC26	*MATα ho::LYS2 lys2Δ ura3Δ arg4 leu2 msh2Δ::KanMX (SK1)*	Crawford et al, 2018
MC49	*MATα; ade8Δ msh2Δ::KanMX (S288c)*	Crawford et al, 2018
MC30	*MATα ho::LYS2 lys2Δ ura3Δ arg4 leu2 msh2Δ::KanMX tel1Δ::HphMX4 (SK1)*	This study
MC53	*MATα; ade8Δ msh2Δ::KanMX tel1Δ::HphMX4 (S288c)*	This study
BLY727	*MATα, msh2::HPH, mlh1::KanMX6 (SK1)*	Marsolier-Kergoat et al 2018
BLY723	*MATα, msh2::HPH, mlh1::KanMX6 (S288c)*	Marsolier-Kergoat et al 2018
BLY372	*MATα, msh2::HPH, mlh3::KanMX6 (SK1)*	Marsolier-Kergoat et al 2018
BLY365	*MATα, msh2::HPH, mlh3::KanMX6 (S288c)*	Marsolier-Kergoat et al 2018
BLY912	*MATα, msh2::HPH, exo1::KanMX4 (SK1)*	Marsolier-Kergoat et al 2018
BLY1070	*MATα, msh2::HPH, exo1::KanMX4 (S288c)*	Marsolier-Kergoat et al 2018

Acknowledgements

We thank Shintaro Yamada and Neeman Mohibullah for help accessing and analysing the mouse and yeast Spo11-oligo datasets, Marie-Claude Marsolier-Kergoat for sharing Python scripts, Jesús Carballo and Michael Lichten for sharing *S. cerevisiae* strains containing relevant constructs (*tel1*Δ::hphNT2 and *sae2*Δ::kanMX6, respectively), Keith Caldecott and Antony Oliver for sharing recombinant TDP2, and Rachal Allison for critical reading of the manuscript.

Funding statement

D.J., V.G., T.J.C., and M.J.N. were supported by an ERC Consolidator Grant (311336), the BBSRC (BB/M010279/1), the Wellcome Trust (200843/Z/16/Z), and a Career Development Award from the Human Frontier Science Program (CDA00060/2010). B.L. and V.G. were supported by the ANR-13-BSV6-0012-01 and ANR-16-CE12-0028-01 grants from the Agence Nationale de la Recherche and a grant from the Fondation ARC pour la Recherche sur le Cancer (PJA20181207756). Work in the S.K. lab was supported by the Howard Hughes Medical Institute; MSK core facilities are supported by National Institutes of Health grant P30 CA008748.

Data availability

Raw *S. cerevisiae* and mouse Spo11-oligo FASTQ data were obtained from published archives GSE84896 and GSE84689 respectively. Nucleotide-resolution maps generated by paired-end Bowtie2 alignment are provided as supplementary files. For mouse, maps used here were generated from the following biological samples: wild type, GSM2247728; *Atm*$^{-/-}$, GSM2247731. FASTQ files used for mapping HR patterns in *S. cerevisiae* SK1 × S288c F1 hybrid octads in *msh2*Δ, *tel1*Δ *msh2*Δ, and *mlh1*Δ *msh2*Δ, *mlh3*Δ *msh2*Δ, *exo1*Δ *msh2*Δ are deposited in the following NCBI SRA archives, PRJNA479661, PRJNA480956, and PRJNA393087[19] respectively. Additional data files are included in the Source Data File, and Supplementary Tables.

References

1. Keeney S, Giroux CN, Kleckner N. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell. 1997; 88 :375–384. [PubMed: 9039264]
2. Bergerat A, et al. An atypical topoisomerase II from Archaea with implications for meiotic recombination. Nature. 1997; 386 :414–417. [PubMed: 9121560]
3. Pan J, et al. A Hierarchical Combination of Factors Shapes the Genome-wide Topography of Yeast Meiotic Recombination Initiation. Cell. 2011; 144 :719–731. [PubMed: 21376234]
4. Lange J, et al. The Landscape of Mouse Meiotic Double-Strand Break Formation, Processing, and Repair. Cell. 2016
5. Neale MJ, Pan J, Keeney S. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. Nature. 2005; 436 :1053–1057. [PubMed: 16107854]
6. Garcia V, Phelps SEL, Gray S, Neale MJ. Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. Nature. 2011; 479 :241–244. [PubMed: 22002605]
7. Fowler KR, Sasaki M, Milman N, Keeney S, Smith GR. Evolutionarily diverse determinants of meiotic DNA break and recombination landscapes across the genome. Genome Res. 2014; 24 :1650–1664. [PubMed: 25024163]
8. Choi K, et al. Nucleosomes and DNA methylation shape meiotic DSB frequency inArabidopsis thalianatransposons and gene regulatory regions. Genome Res. 2018
9. Cannavo E, Cejka P. Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. Nature. 2014; 514 :122–125. [PubMed: 25231868]
10. Garcia V, Gray S, Allison RM, Cooper TJ, Neale MJ. Tel1(ATM)-mediated interference suppresses clustered meiotic double-strand-break formation. Nature. 2015; 520 :114–118. [PubMed: 25539084]

11. Mohibullah N, Keeney S. Numerical and spatial patterning of yeast meiotic DNA breaks by Tel1. Genome Res. 2017; 27 :278–288. [PubMed: 27923845]
12. Zhang L, Kleckner NE, Storlazzi A, Kim KP. Meiotic double-strand breaks occur once per pair of (sister) chromatids and, via Mec1IATR and Tel1IATM, once per quartet of chromatids. Proc Natl Acad Sci U S A. 2011; 108 :20036–20041. [PubMed: 22123968]
13. Joyce EF, et al. Drosophila ATM and ATR have distinct activities in the regulation of meiotic DNA damage and repair. J Cell Biol. 2011; 195 :359–367. [PubMed: 22024169]
14. Lange J, et al. ATM controls meiotic double-strand-break formation. Nature. 2011; 479 :237–240. [PubMed: 22002603]
15. Carballo JA, et al. Budding Yeast ATM/ATR Control Meiotic Double-Strand Break (DSB) Levels by Down-Regulating Rec114, an Essential Component of the DSB-machinery. PLoS Genet. 2013; 9 e1003545 [PubMed: 23825959]
16. Liu J, Wu TC, Lichten M. The location and structure of double-strand DNA breaks induced during yeast meiosis: evidence for a covalently linked DNA-protein intermediate. EMBO J. 1995; 14 :4599–4608. [PubMed: 7556103]
17. Claeys Bouuaert C, et al. Structural and functional characterization of the Spo11 core complex. Nat Struct Mol Biol. 2021; 28 :92–102. [PubMed: 33398171]
18. Martini E, et al. Genome-wide analysis of heteroduplex DNA in mismatch repair-deficient yeast cells reveals novel properties of meiotic recombination pathways. PLoS Genet. 2011; 7 e1002305 [PubMed: 21980306]
19. Marsolier-Kergoat MC, Khan MM, Schott J, Zhu X, Llorente B. Mechanistic View and Genetic Control of DNA Recombination during Meiosis. Mol Cell. 2018; 70 :9–20. e6 [PubMed: 29625041]
20. Kulkarni DS, et al. PCNA activates the MutLy endonuclease to promote meiotic crossing over. Nature. 2020; 586 :623–627. [PubMed: 32814343]
21. Cannavo E, et al. Regulation of the MLH1-MLH3 endonuclease in meiosis. Nature. 2020; 586 :618–622. [PubMed: 32814904]
22. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. Cell. 1983; 33 :25–35. [PubMed: 6380756]
23. Diaz RL, Alcid AD, Berger JM, Keeney S. Identification of residues in yeast Spo11 p critical for meiotic DNA double-strand break formation. Mol Cell Biol. 2002; 22 :1106–1115. [PubMed: 11809802]
24. Noll M. Nucleic Acids Research. 1974; 1 :1573–1578. [PubMed: 10793712]
25. Gittens WH, et al. A nucleotide resolution map of Top2-linked DNA breaks in the yeast and human genome. Nature Communications. 2019; 10
26. Prieler S, Penkner A, Borde V, Klein F. The control of Spo11‘s interaction with meiotic recombination hotspots. Genes Dev. 2005; 19 :255–269. [PubMed: 15655113]
27. Panizza S, et al. Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. Cell. 2011; 146 :372–383. [PubMed: 21816273]
28. Kugou K, et al. Rec8 guides canonical Spo11 distribution along yeast meiotic chromosomes. Mol Biol Cell. 2009; 20 :3064–3076. [PubMed: 19439448]
29. Li J, Hooker GW, Roeder GS. Saccharomyces cerevisiae Mer2, Mei4 and Rec114 form a complex required for meiotic double-strand break formation. Genetics. 2006; 173 :1969–1981. [PubMed: 16783010]
30. Kee K, Protacio RU, Arora C, Keeney S. Spatial organization and dynamics of the association of Rec102 and Rec104 with meiotic chromosomes. EMBO J. 2004; 23 :1815–1824. [PubMed: 15044957]
31. Kumar R, Bourbon HM, de Massy B. Functional conservation of Mei4 for meiotic DNA double-strand break formation from yeasts to mice. Genes Dev. 2010; 24 :1266–1280. [PubMed: 20551173]
32. Claeys Bouuaert C, Pu S, Wang J, Patel DJ, Keeney S. DNA-dependent macromolecular condensation drives self-assembly of the meiotic DNA break machinery. bioRxiv. 2020; doi: 10.1101/2020.02.21.960245

33. Lukaszewicz A, Lange J, Keeney S, Jasin M. De novo deletion mutations at recombination hotspots in mouse germlines. bioRxiv. 2020; doi: 10.1101/2020.06.23.168138v1
34. Crawford M, Cooper TJ, Marsolier-Kergoat M-C, Llorente B, Neale MJ. Separable roles of the DNA damage response kinase Mec1(ATR) and its activator Rad24(RAD17) within the regulation of meiotic recombination. bioRxiv. 2018
35. Johnson D, Allison RM, Cannavo E, Cejka P, Neale M. Removal of Spo11 from meiotic DNA breaks in vitro but not in vivo by Tyrosyl DNA Phosphodiesterase 2. bioRxiv. 2019; doi: 10.1101/527333
36. Xu L, Kleckner N. Sequence non-specific double-strand breaks and interhomolog interactions prior to double-strand break formation at a meiotic recombination hot spot in yeast. EMBO J. 1995; 14 :5115–5128. [PubMed: 7588640]
37. Cortes Ledesma F, El Khamisy SF, Zuma MC, Osborn K, Caldecott KW. A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. Nature. 2009; 461 :674–678. [PubMed: 19794497]
38. Cao L, Alani E, Kleckner N. A pathway for generation and processing of double-strand breaks during meiotic recombination in S. cerevisiae. Cell. 1990; 61 :1089–1101. [PubMed: 2190690]
39. Alani E, Padmore R, Kleckner N. Analysis of wild-type and rad50 mutants of yeast suggests an intimate relationship between meiotic chromosome synapsis and recombination. Cell. 1990; 61 :419–436. [PubMed: 2185891]
40. Furuse M, et al. Distinct roles of two separable in vitro activities of yeast Mre11 in mitotic and meiotic recombination. EMBO J. 1998; 17 :6412–6425. [PubMed: 9799249]
41. Moreau S, Ferguson JR, Symington LS. The nuclease activity of Mre11 is required for meiosis but not for mating type switching, end joining, or telomere maintenance. Mol Cell Biol. 1999; 19 :556–566. [PubMed: 9858579]
42. Gittens W, et al. A nucleotide resolution map of Top2-linked DNA breaks in the yeast and human genome. bioRxiv. 2019
43. Blat Y, Protacio RU, Hunter N, Kleckner N. Physical and functional interactions among basic chromosome organizational features govern early steps of meiotic chiasma formation. Cell. 2002; 111 :791–802. [PubMed: 12526806]
44. Acquaviva L, et al. The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. Science. 2013; 339 :215–218. [PubMed: 23160953]
45. Sommermeyer V, Béneut C, Chaplais E, Serrentino ME, Borde V. Spp1, a Member of the Set1 Complex, Promotes Meiotic DSB Formation in Promoters by Tethering Histone H3K4 Methylation Sites to Chromosome Axes. Mol Cell. 2013; 49 :43–54. [PubMed: 23246437]
46. Kane SM, Roth R. Carbohydrate metabolism during ascospore development in yeast. J Bacteriol. 1974; 118 :8–14. [PubMed: 4595206]

Europe PMC Funders Author Manuscripts

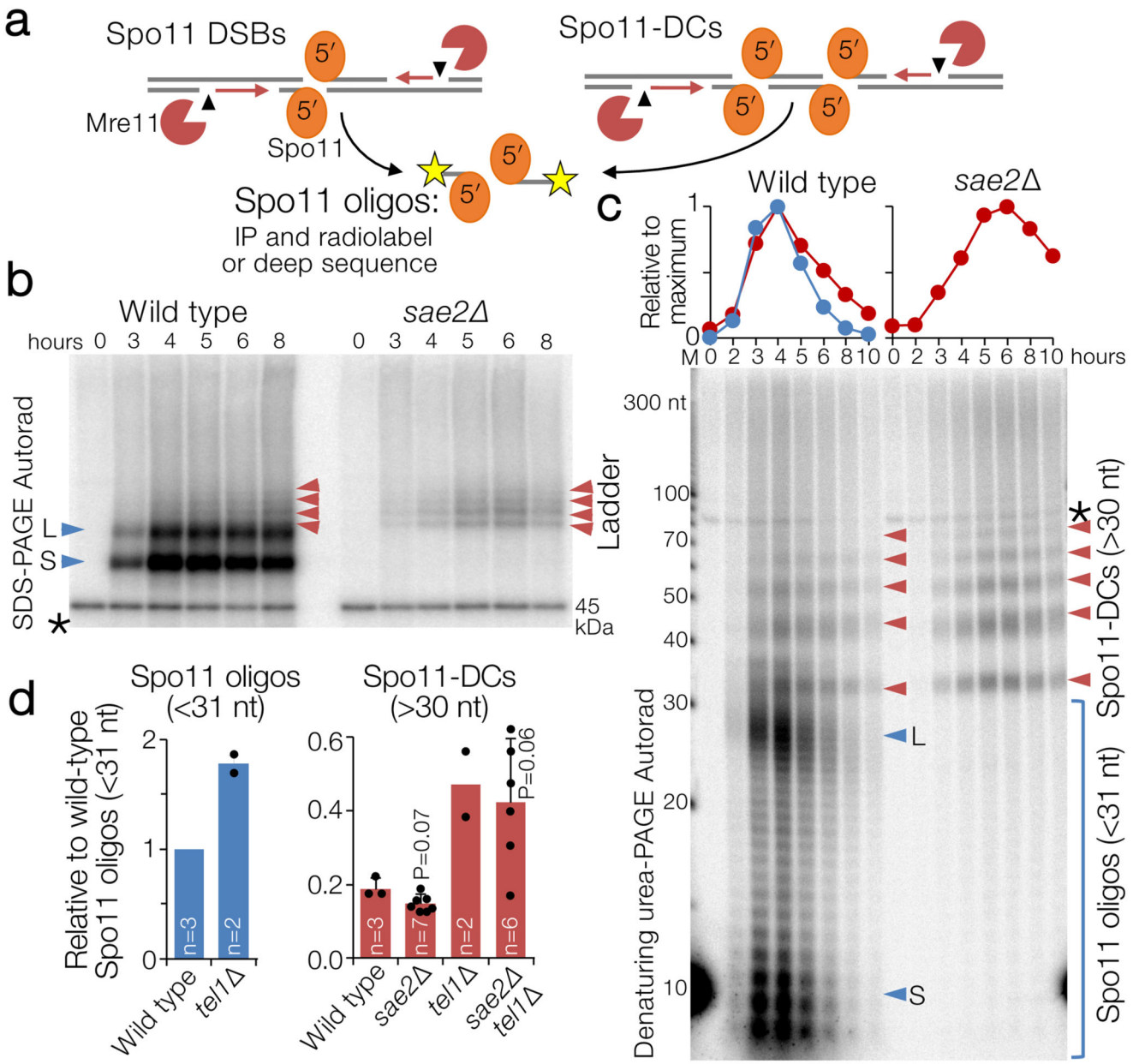


Figure 1. Spo11-DC formation during *S. cerevisiae* meiosis.

a, Spo11 DSBs, processed by Mre11 endo- (arrowhead) and exo- (red arrows) nucleolytic activities, generate canonical Spo11 oligos. Adjacent Spo11 DSBs ("double cutting", Spo11-DC) can create a distinct class of Spo11 oligos independently of Mre11. **b-c**, Immunoprecipitated (IP) FLAG-tagged Spo11-oligo and Spo11-DC products separated by 8% SDS-PAGE (**b**) or 19% denaturing PAGE following treatment with Proteinase K and TDP2 (**c**) at hourly timepoints during meiosis (see also Extended Data Fig.1). Red arrows: Spo11-DC ladder; blue arrows: long (L) and short (S) canonical Spo11 oligos. Spo11 oligos and 10 nt marker (M) were radiolabelled with chain-terminating [α-^{32}P]-3′-dATP using terminal transferase. Asterisk: labelling artefact. **d**, Quantification of Spo11 oligos and

Spo11-DC signals at peak timepoints (4-6 h) detected as in (**c**) relative to wild-type Spo11 oligos (n=biologically independent samples, mean±SD is indicated; P=unpaired two-tailed T test).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

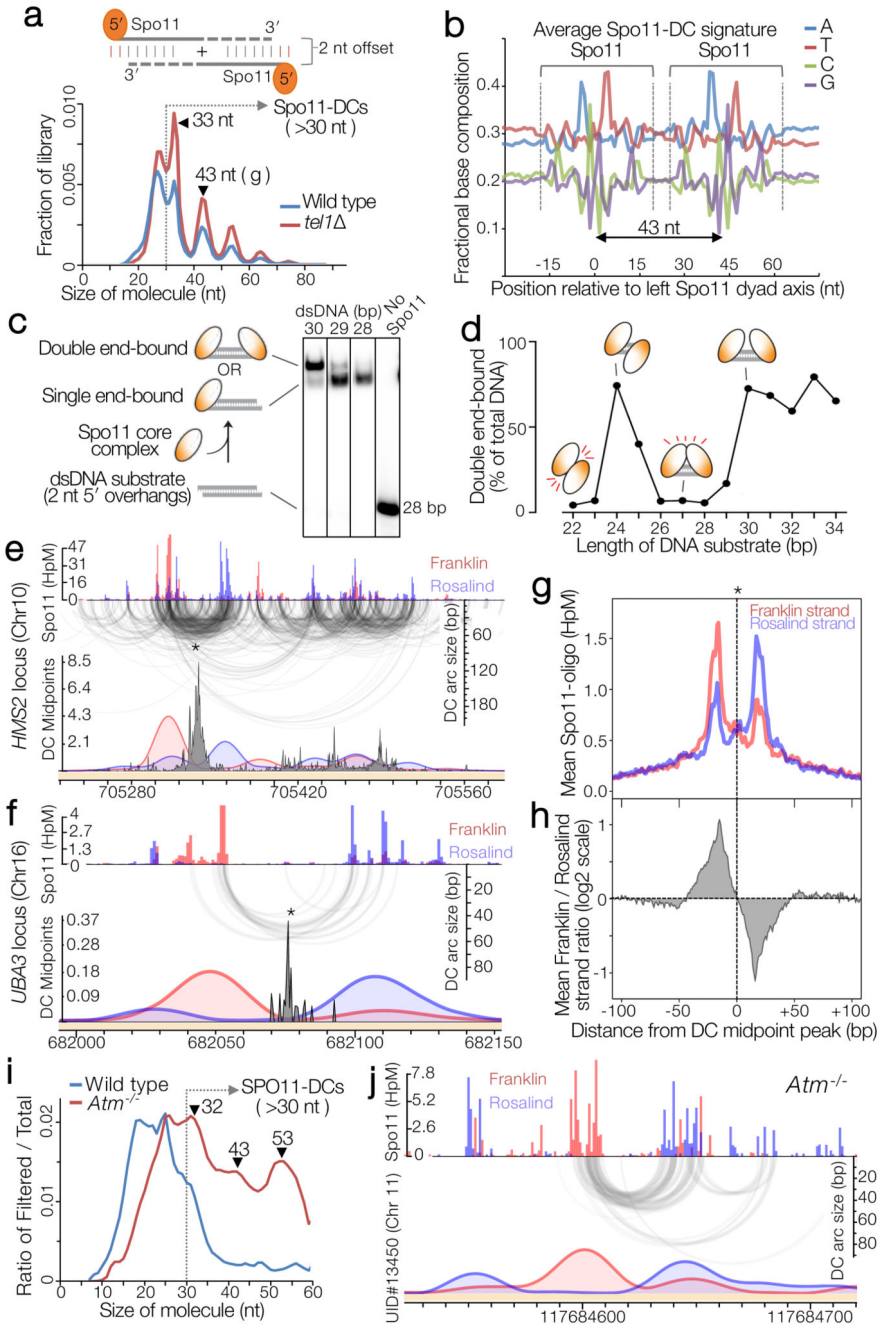


Figure 2. Spo11-DC identification within Spo11-oligo libraries.

a, Size distribution of *S. cerevisiae* Spo11 oligos[11] after filtering to retain only overlapping pairs of molecules with a 2-nt 5′ overhang at both ends (Spo11-DCs). **b,** Nucleotide composition of 43-nt filtered molecules. **c,** *In vitro* double-end binding assay. Spo11-Rec102-Rec104-Ski8 'core complex' was incubated with double-stranded DNA substrates with 2-nt 5′ overhangs, mimicking the product of Spo11-mediated cleavage. The core complex binds such DNA ends with high affinity[17]. **d**, Quantification of double-end bound complexes with DNA substrates of different lengths (Extended Data Fig.2j). **e-f**, Arc

diagram of Spo11-DCs (Extended Data Fig.4a) mapped within two example hotspots. Top panel: unfiltered strand-specific Spo11 oligos (Franklin-strand, red; Rosalind-strand, blue; HpM, hits per million mapped reads). Arcs (grey-scale frequency-weighted) link 5′ ends of each Spo11-DC. Lower panel: smoothed unfiltered strand-specific Spo11 oligos, overlaid with frequency histogram of Spo11-DC midpoints (grey). Note the spatial association of Spo11-DC sub-domains with asymmetric Franklin-Rosalind strand disparity. **g-h,** Strand disparity at Spo11-oligo 5′ ends is a genome-wide feature of Spo11-DCs (see also Extended Data Fig.5). Average strand-specific Spo11-oligo signal (**g**), and strand ratio (**h**), centred upon the strongest Spo11-DC midpoint (asterisks in **e-f**) within every annotated DSB hotspot (3910 loci). **i-j,** Fine-scale analysis of SPO11-DCs within mouse hotspots (see also Extended Data Fig.6). Filtered SPO11-oligo[4] length distribution enriches peaks in *Atm*$^{-/-}$ that display ~10 nt periodicity **(i)**. Arc diagrams of SPO11-DCs in *Atm*$^{-/-}$ (as in **e-f**).

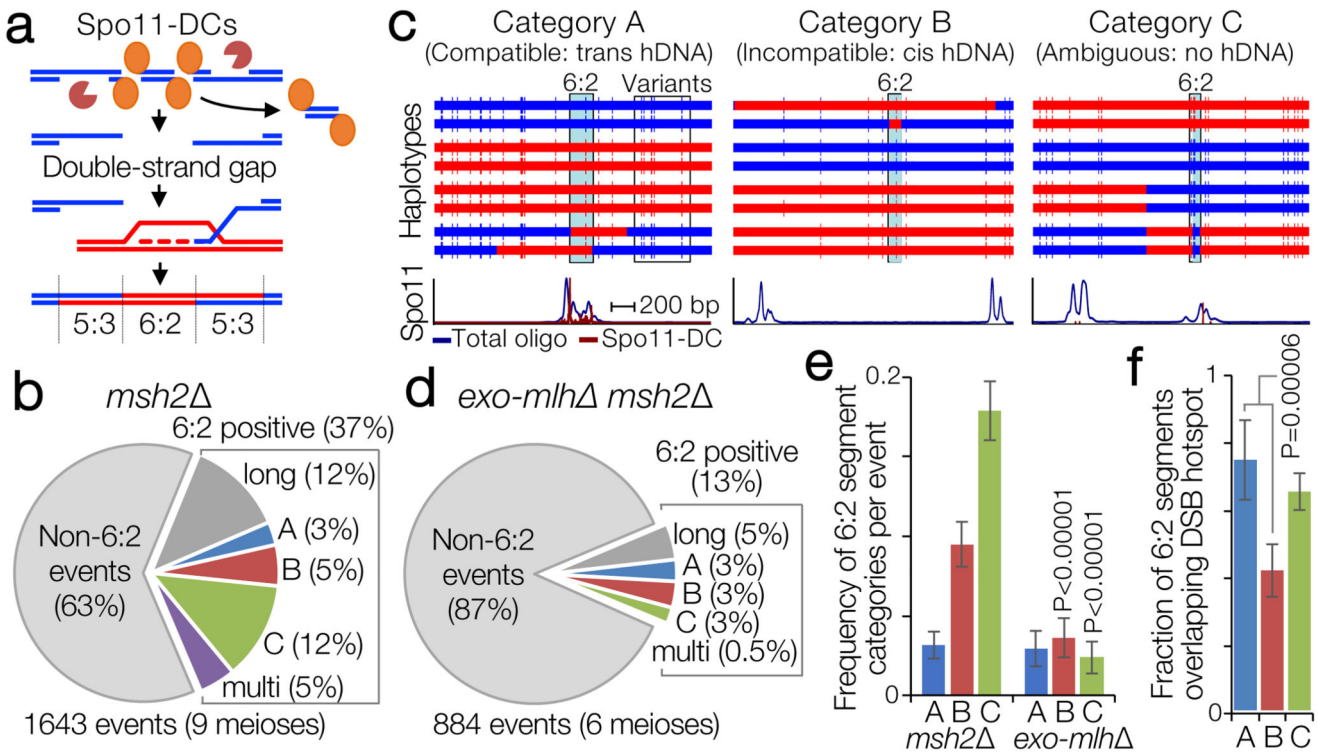


Figure 3. Spo11-DC-induced gap repair in *S. cerevisiae*.

a, In mismatch repair-deficient strains *(msh2Δ)* gap repair may generate segments of 6:2 marker segregation flanked by 'trans' 5:3 hDNA. **b**, Quantification of meiotic recombination event types in control (*msh2*Δ). **c**, Representative categorisations of 6:2-containing events based on flanking hDNA (see also Extended Data Fig.7). Horizontal bars are sequence polymorphism calls (red or blue) from meiotic octads. 6:2 segments indicated in pale blue. Lower panels are smoothed histograms of total Spo11-oligo and Spo11-DC signals. **d**, As in (**b**), but for pooled *'exo-mlhΔ msh2Δ' (exo1Δ msh2Δ, mlh1Δ msh2Δ, mlh3Δ msh2Δ)* strains. **e**, Frequency of 6:2 segments per event for each category. **f**, Fraction of 6:2 segments (in *msh2Δ*) overlapping annotated hotspots. P values indicate two-tailed Z-test of proportions relative to *msh2Δ* (**e**), or between categories (**f**). In (**e-f**) whiskers are 95% confidence intervals; n=9 *(msh2Δ)* or 6 *(exo-mlhΔ)* biologically independent samples.

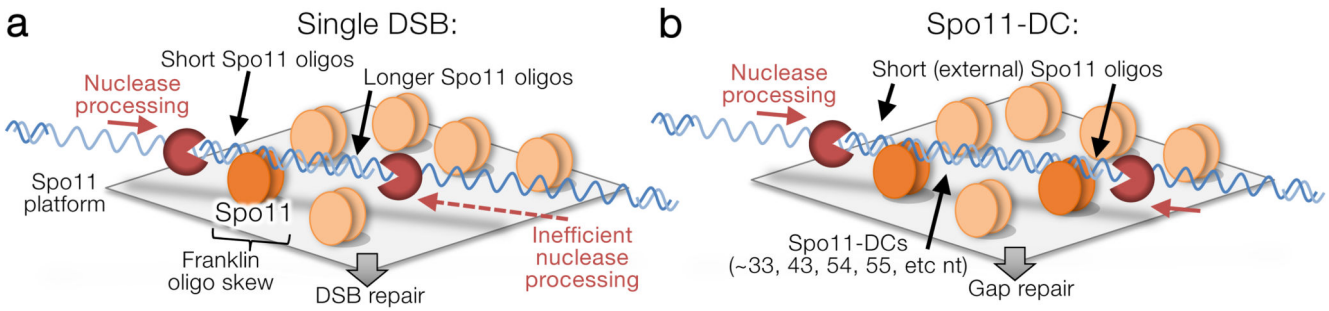


Figure 4. Model for Spo11-DC formation.

a-b, Hotspot DNA interacts with multiple Spo11 dimers (pale orange ellipses) assembled within a surface of pro-DSB axis factors. Single (**a**) or concerted (**b**) DSBs may form depending upon favoured interactions between Spo11 (dark orange Spo11 dimers) and the repeating structure of the DNA helix. Concerted DSB formation (**b**) generates internal Spo11-DCs with lengths that are multiples of the helical pitch (~10.5 bp). Our model is agnostic as to whether Spo11 is organised with any repeating pattern. Rather, periodic Spo11-DC sizes arise from the relative stiffness of B-form DNA and the constrained interaction of Spo11 with only one face of the DNA helix. Resection initiates efficiently in flanking regions, generating short external Spo11 oligos and initiating gap repair. Single DSBs (**a**) are processed by nucleases on both sides, but resection proceeds less efficiently within the axis-associated DNA, generating a Spo11-oligo length asymmetry and consequent mapped strand disparity (see Extended Data Fig.8a).

Europe PMC Funders Author Manuscripts