The EuroBiotech Journal
RESEARCH ARTICLE
Medical Genetics

cis- and *trans*- regulation controls of human meiotic recombination at a hotspot

Mahmut C. Ergören[1,2], Rita Neumann[2], Ingrid Berg[2] and Alec J. Jeffreys[2]

Abstract

PRDM9 plays a key role in specifying meiotic recombination hotspot locations in humans. To examine the effects of both the 13-bp sequence motif (cis-regulator) and trans-regulator PRDM9 on crossover frequencies and distribution, we studied Hotspot DA. This hotspot had the motif at its centre, and a single nucleotide polymorphism (SNP) that disrupts the motif. The crossover frequency showed Hotspot DA to be a regular hotspot with an average crossover rate (~8 X10-4) among hotspots assayed on autosomes. Our results show that, comparing the rates and distributions of sperm crossover events between donors heterozygous for the disrupting SNP showed that there was a huge asymmetry between the two alleles, with the derived, motif-disrupting allele completely suppressing hotspot activity. Intensive biased gene conversion, both in to crossovers and noncrossovers, has been found at Hotspot DA. Biased gene conversion that influences crossover and non-crossover hotspot activity correlates with PRDM9 allele A. In Hotspot DA, the lifetime of the hotspot mostly depends on the cis-regulatory disrupting SNP, and on the trans-regulatory factor PRDM9. Overall, our observation showed that Hotspot DA is the only evidence for human crossover hotspot regulation by a very strong cisregulatory disrupting SNP.

[1]Depatment of Medical Biology, Faculty of Medicine, Near East University, Nicosia, Mersin, Turkey

[2]Department of Genetics, Faculty of Biological Sciences and Medicine, University of Leicester, Leicester, United Kingdom

Corresponding author: M. C. Ergören
E-mail: mahmutcerkez@gmail.com

Published online: 27 October 2017
doi:10.24190/ISSN2564-615X/2017/04.09

Introduction

Meiotic recombination is an important evolutionary force in shaping the human genome and creating human diversity. In humans, at least one recombination event per chromosome is required for proper segregation of chromosomes, and in cases of non-disjunction, genetic disorders such as Down's syndrome may result. The recombination event may be resolved as a crossover, where information is reciprocally exchanged between chromosomes, or a non-exchange gene conversion event. Recombination events have mostly been studied in yeast and mice. The reciprocal exchange that takes place during meiotic division can occur nearly anywhere along a chromosome. However, meiotic recombination occurs more frequently in some regions of the genome called hotspots, which arise in 1-2 kb intervals across the human genome (1).

Furthermore, our knowledge about regulating factors of the human recombination machinery is very limited. Sperm typing has been used most successfully to identify hotspots in humans. The first characterized hotspot using sperm typing was the MS32 hotspot on chromosome 1 (1). Biased gene conversion had been found responsible for over-transmitting alleles, and therefore implicated in hotspot silencing (2). Recombination hotspots could be influenced by single base changes known as single nucleotide polymorphisms (SNPs) (3). Genome-wide comparison of Linkage Disequilibrium (LD) hotspots has identified 13-bp CCNCCNTNNCCNC DNA sequence motif associated with 40% of these hotspots and it is thought to play an important role in recombination (4). Myers *et al.,* (2005) has identified a few LD hotspots candidates from low-resolution HapMap Phase II data, moreover, these LD hotspots based on having a 13-bp CCNCCNTNNCCNC motif with a disrupting SNP that thought to disrupt the motif and can down-regulate recombination activity (4, 5).

Recently, the meiosis-specific protein PR domain–containing 9, PRDM9, has been identified as a major regulator of hotspots in the human and mouse genome (6, 8). PRDM9 is a histone H3 methyltransferase with a C-terminal tandem-repeat C2H2 zinc finger (ZnF) domain encoded by a minisatellite (9). It is uniquely expressed during early meiosis in both males and females (8).

Recent studies revealed that, depending on the carried *PRDM9* alleles, the recombination profiles and crossover hotspot activity showed variation in mouse subspecies hybrids (8). In humans, according to computational analyses (7), PRDM9 has been identified as a human ZnF protein that recognizes the 13-bp CCNCCNTNNCCNC motif. This motif might also serve as a binding site for the ZnF protein (4, 7) Nevertheless, an allele of *PRDM9* is the most common allele that binds *in vitro*, but the *PRDM9* allele 'I' cannot bind to this motif (6).

Previously Berg *et al.,* (2010) investigated the influence of variation in the PRDM9 ZnF array and the results demonstrated that crossover activity at individual human recombination hotspots and genome instability both at minisatellites and at pathological non-allelic homologous recombination (NAHR) rearrangements, are all influenced by PRDM9 variation. Sixteen different forms of PRDM9 containing between 8 and 18 zinc fingers have been found in 74 African and 156 European semen donors (10). Ten active hotspots (including 5 hotspots with the sequence motif) were examined, and the results showed that all 10 hotspots showed activity dependent on PRDM9, and were generally activated by the common allele 'A'. This study implies that PRDM9 operates across a much longer track than the 13-bp sequence motif (10). The most recent study shows that the activation of PRDM9 variants that are common in Africans but rare in Europeans reveals second-class hotspots, and these hotspots act differently even in the same populations (11).

Hereby, we firstly confirmed the presence of a LD hotspot (Hotspot DA) and whether the disrupting SNP (rs7036542 G>T, DA7.5G/T) within the motif located at the centre of the hotspot by high-resolution genotyping techniques used for genotype out European semen donor panel. Secondly, the analysis of recombination rate and distribution, and determination of the hotspot centre of Hotspot DA carried out by LDU analysis. The crossover frequencies and distributions were compared with previously established parameters (12). Also, biased gene conversion was determined by comparing recombination frequencies and distributions in both reciprocal orientations. Interestingly, the results show the highest observed transmission distortion ratio. Furthermore, it was tested whether crossovers and non-crossovers are influenced by the same biases. For better understanding of the effect of *cis*-regulation (rs7036542) on the studied hotspot, the analysis was expanded to more men to allow the identification of DNA sequence motifs that carry the disrupting SNP to control hotspot activation. Hotspot DA shows for the first time an example of strong *cis*-regulatory control on a human crossover hotspot. Finally, we answered the important remain question, what is the effect of the *trans*-regulator factor PRDM9 on the activity of Hotspot DA? To answer this question, more men with variant PRDM9 genotypes were analyzed to explore possible influences of PRDM9 status on crossover frequencies. To conclude, on Hotspot DA we showed for the first time whether *cis*-regulation (rs7036542) is more effective than the *trans*-regulator factor PRDM9.

Materials and Methods

Ethic statement

Ethics Statement Semen samples were collected with informed consent and approval from the Leicestershire Health Authority Research Ethics Committee (ref 6659).

Sample preparation

Sperm DNA was isolated as described previously (1) and quantified on a NanoDrop1000 spectrophotometer. For details of donors, see (10, 11).

Hotspot selection

The LD hotspot DA is firstly identified by low-resolution HapMap Phase II analyses (5). Hotspot DA contains an intergenic SNP (rs7036542), which is located on the short arm of chromosome 9 in its sub terminal region, and within a well-localized putative hotspot within a THE1B element. This SNP changes the 13-bp motif from CCNCCNTNNCCNC to ACNCCNTNNCCNC. The ancestral SNP allele is C (or complementary G). . LDMAP was used to create metric LD maps (12) across the subtelomeric regions of the chromosome 9 genotype data from the Phase III HapMap dataset (release 27) (http://www.hapmap.org). Historical recombination frequency at this LD hotspot was estimated by coalescent analysis by using LDhat (13), assuming effective population sizes of 10,000.

SNP genotyping

SNP Genotyping Routine genotyping was performed on whole-genome amplified DNA, generated from 40 ng aliquots of each DNA using the GenomiPhi HY DNA amplification kit (GE Healthcare BioSciences). Hotspot target regions were amplified in several partially overlapping PCRs by successive rounds of nested PCR, the products transferred onto nylon membranes and genotyped by allele-specific oligonucleotide (ASO) hybridization as described previously (11, 14, 15). The linkage phase of internal SNPs was determined by testing separated haplotypes by sequential ASO hybridization at each of the heterozygous SNPs. All those universal primers, ASOs, allele-specific primers and PCR conditions (18).

Re-sequencing of the donors

The hotspot centre region was re-sequenced in men chosen for analysis, to identify any additional SNPs. Parental haplotypes of test individuals were separately amplified using allele-specific primers directed to markers outside of the hotspot interval. Excess primer and unincorporated dNTPs were removed using 1.4 U/ml exonuclease I (New England Biolabs) and 0.21 U/ml shrimp alkaline phosphatase (Roche), with incubation at 37 uC for 60 min, followed by 15 min at 80 uC. Universal se-

quencing primers were designed for 3–4 targets covering the hotspot, with targets overlapping by 100 bp to ensure complete coverage. Standard 20 ml Big Dye Terminator v 3.1 sequencing reactions were carried out, the extension products purified using Performa DTR-gel filtration Cartridges and then separated on a 3730 DNA Analyser (Applied Biosystems).

Cross-Over analysis

All 74 African semen donors plus 103 Europeans and two Indians, including all those with atypical-length PRDM9 alleles, were genotyped for all SNPs for the LD Hotspot DA. Men for cross-over analysis were selected based on appropriate hotspot genotypes and PRDM9 status.

Cross-over mapping was conducted as described (15, 16) by using information on cross-over frequencies to optimize the recovery of typically 120 cross-over molecules in each orientation for each man tested. Recombinants were detected by the presence of markers from the unselected haplotype in the amplified DNA. Cross-over and conversion frequencies were estimated by assuming a single DNA molecule PCR efficiency of 50% throughout (15, 16). The frequencies and distributions of

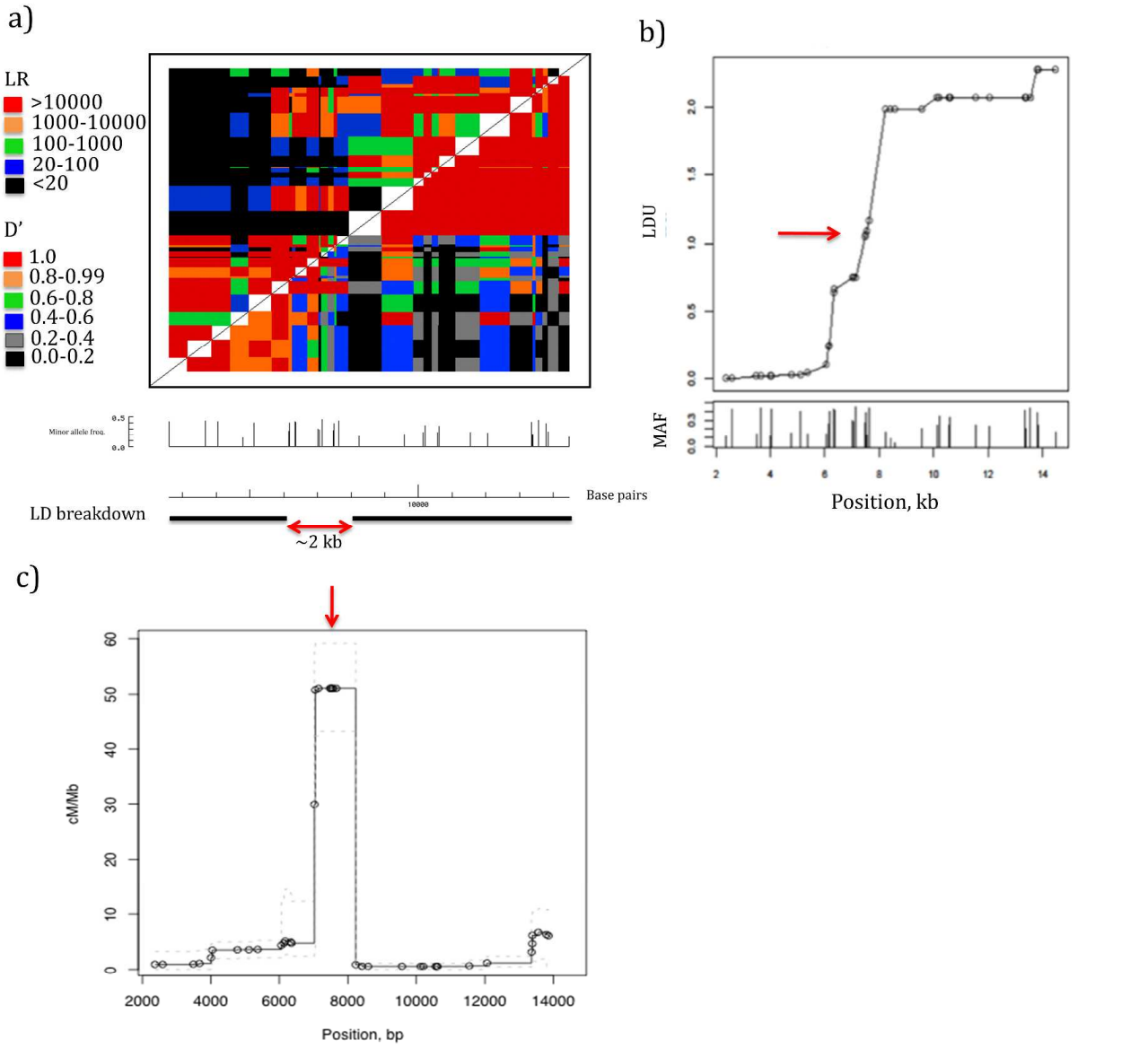

Figure 1. a) Linkage disequilibrium breakdown across the target region DA. All analyses were based on genotyped SNP typed data for Hotspot DA in donor panel. Maximum likelihood haplotype frequencies for each pair of SNPs were determined and used to estimate |D'| levels of LD (lower right), plus the associated likelihood ratio (LR, odds of LD) versus free association (upper left), and are colour-coded as indicated. Only SNPs with minor allele frequencies (MAFs) ≥ 0.15 were included in the analysis. The locations of the SNPs are shown below and to the right of the plot, with positions centred on the middle of DA at co-ordinate 0. The LD block is shown in black below the plot, and the position and the approximate width of the LD hotspot DA is indicated by the red arrow. b) LDU map of Hotspot DA. LDU map confirmed the putative hotspots with 1-2 kb intervals. The disrupting SNP within the motif is shown with red arrow. In Hotpsot DA, the disrupting SNP are localized close to the centre. c) Coalescent analys of a LD hotspot DA. Coalescent mapping gives clues about historical recombination and recombination frequency. Hotspot DA showed a highest peak activity of ~ 50 cM/Mb crossover rate (RF ~ 5 × 10-4). The width of Hotspot DA was ~1.4 kb. Therefore, LDU map was confirmed by coalescent analyse for Hotspot D. Also, red arrows show the location of the disrupting SNP within the motif in the LD hotspot.

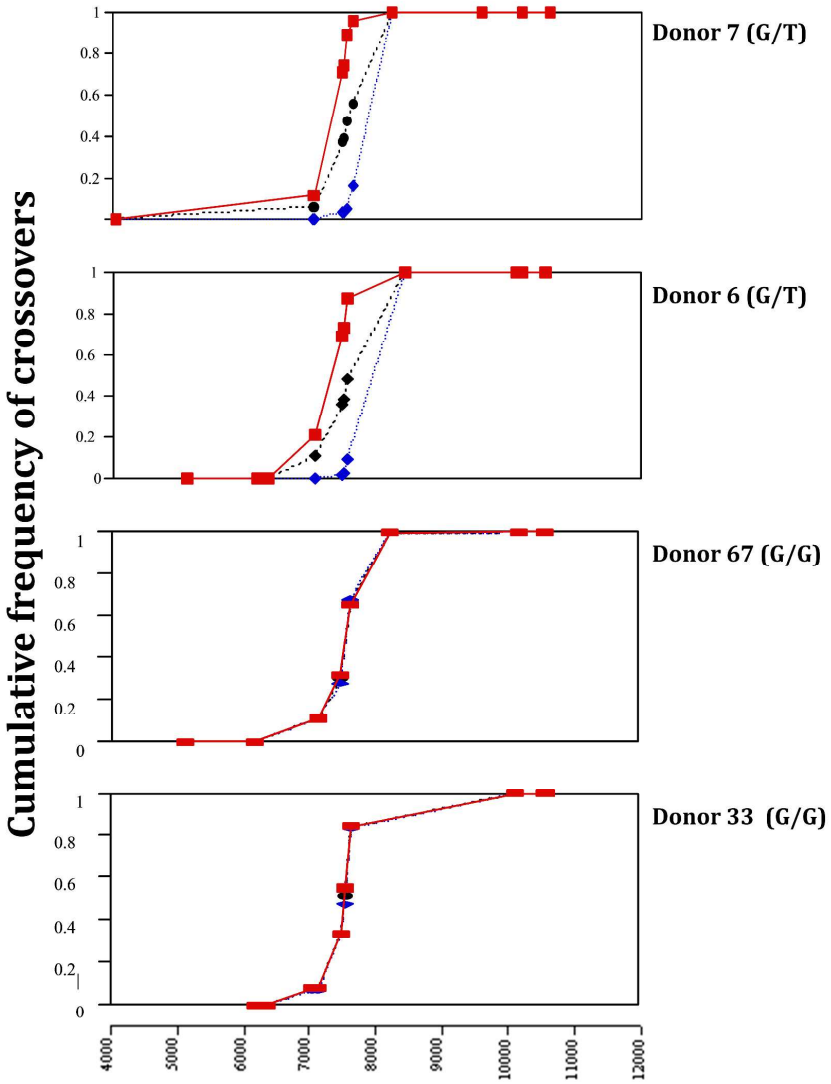

Figure 2. Cumulative frequency of crossovers for Hotspot DA. The cumulative frequency values for orientation A (red), and B (blue) for each man are shown. The best fit curve for orientations A+B (black) is shown for each man. The curves for all four donor show a similar shape. Cumulative distributions shifted by 330-450 bp for A and B crossovers for donor 7 and donor 6. The crossover asymmetry was not observed at donor 67 and donor 33.

cross-overs detected in normal cross-over assays performed on the same men were indistinguishable. A 1.5-kb interval spanning the center of each hotspot was re-sequenced in most mapped donors to maximize the number of SNP markers and to search for variants that might influence hotspot activity in cis.

Results

Location and morphology of the LD Hotspot DA

According to our knowledge, 40% of human recombination hotspots have a 13-bp CCNCCNTNNCCNC sequence motif that determines activity in the hotspots (4). The work of previous studies has shown that CCTCCCT and CCCCACCCC motifs are present at the centre of hotspots *DNA2* and *NIDI* (3, 17). Furthermore, SNPs in both the two motifs disrupt hotspot activity, with the non-reference SNP base acting as a recombination-suppressing allele (5). The SNP, within the 13-bp CCNCCNTNNCCNC motif that is located to the middle of a hotspot, is analyzed to confirm the presence of the putative hotspot in our donor panel by LDU mapping. Confirming this hotspot with LDU mapping was crucial because it establishes the existence of the hotspot in our donor panel and the location of disrupting SNPs within the motif, and also the size of the hotspot can be estimated.

A 15-kb interval of DNA sequence spanning Hotspot DA was downloaded from ENSEMBL (www.ensembl.org). The sequence was annotated to include all SNPs (including both European and African) and repeat sequences (LINEs, Alus etc) using information gained from Phase II HapMap data, dbSNP data and Repeat Masker. Thirty-nine SNPs have been genotyped using ASO hybridisation for the putative hotspot DA (For genotype data, (18)). LDU mapping confirmed 1-2 kb intervals (1) with the disrupting SNP in the middle, and estimated the historical recombination frequency (RF) to be ~5×10^{-4}. This RF shows that the hotspot is an ordinary hotspot (**Fig. 1a** and **1b**). Moreover, coalescent mapping that gives clues about

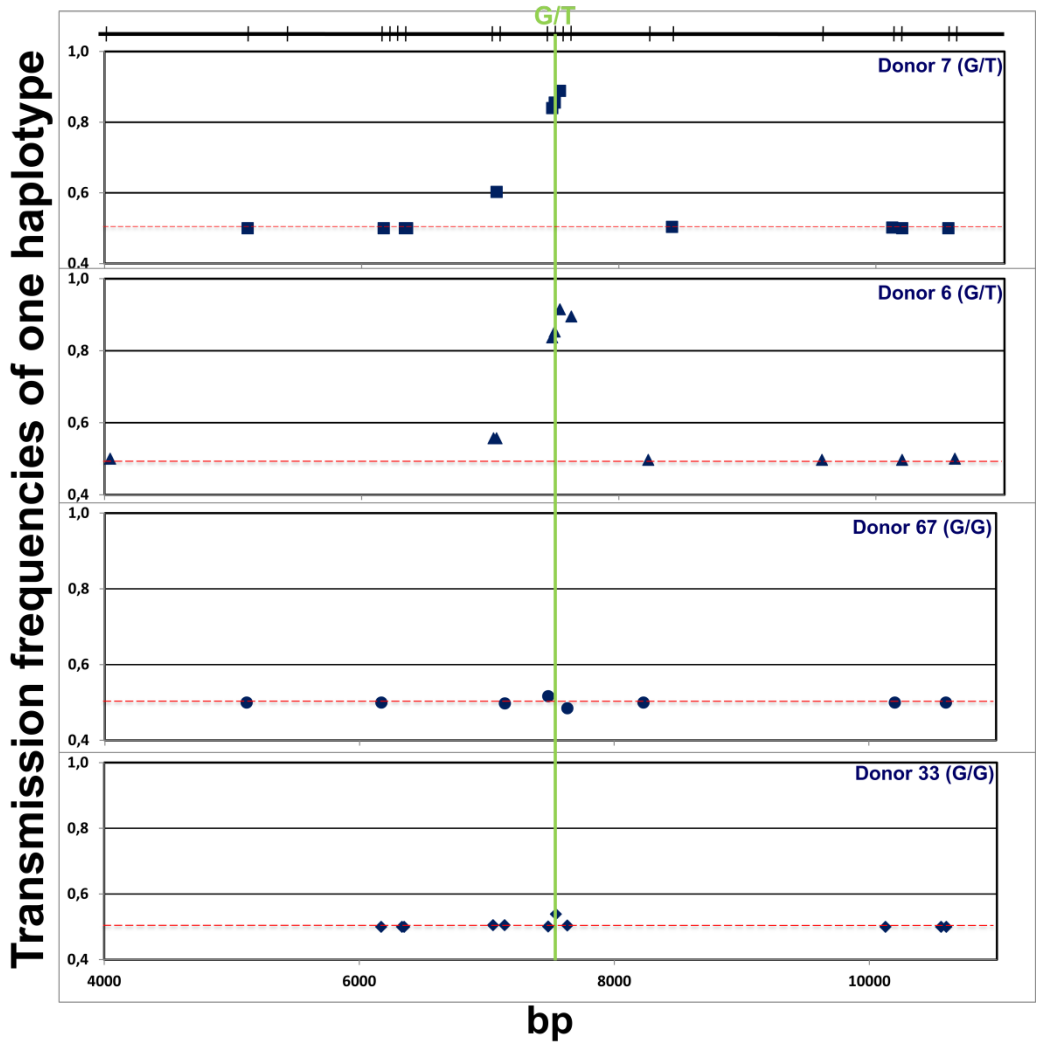

Figure 3. Transmission of alleles from one haplotype into crossover progeny. Central Markers (DA7.5G/T and DA7.5aC/T) showed 90:10 transmission ratios into crossover progeny in donor 7 and donor 6. This biased gene conversion is maximal for the motif-disrupting SNP located just 90 bp from the hotspot centre. However, transmission distortion in donor 67 and donor 33 is much less marked, with central markers showing at most respectively 53:47 and 51:49 transmission ratios. The location of the disrupting SNP DA7.5G/T is shown by the green line.

historical recombination and recombination frequency, carried out for Hotspot DA and it gave the highest peak activity of ~ 50 cM/Mb crossover rate (RF ~ 5×10^{-4}) and the width of Hotspot DA was ~1.4 kb (**Fig. 1c**).

Metric LD analysis (LDU) was performed using 74 African semen donors genotype data (see 18) to see if the African population showed a historical hotspot as well as the Europeans. The LDU map across Hotspot DA showed an increase of LD across the target interval in the African population, implying that historical recombination has taken place. The LDU step of the African population was reduced two-fold compared to the European population (18). However, historical recombination has not necessarily influenced LD to a similar extent in both populations.

Crossover frequency polymorphism and reciprocal crossover asymmetry

A 15-kb target interval around Hotspot DA that was centered on the motif-disrupting SNP 7.5 (rs703654) within the 13-bp hotspot sequence motif was assayed for recombination activity in a total of six men. The crossover rates showed Hotspot DA to be a regular hotspot with an average crossover rate among hotspots assayed on autosomal chromosomes. Hotspot DA provides strong evidence that the motif is likely involved in promoting the initiation of recombination. Four assayed men (donors 7, 6, 67, 33) showed very similar hotspot locations for orientation A + B crossovers combined (**Fig. 2**). This centre is displaced 3′ to DA7.5 by about 50-70 bp. The numbers of crossovers mapping 5′ and 3′ of DA7.5 were normalised to equal numbers of A and B crossovers, and a Fisher test was performed to check if the numbers were significantly different from the 50:50 ratio predicted if DA.7.5 is located exactly at the centre. Donors 6, 7, 67 and 33 showed statistically significant crossovers mapping of DA7.5 was; the only exception was donor 55 where the shift was of borderline significance (P = 0.054, two-tailed Fisher exact test). Thus, DA7.5 and the motif

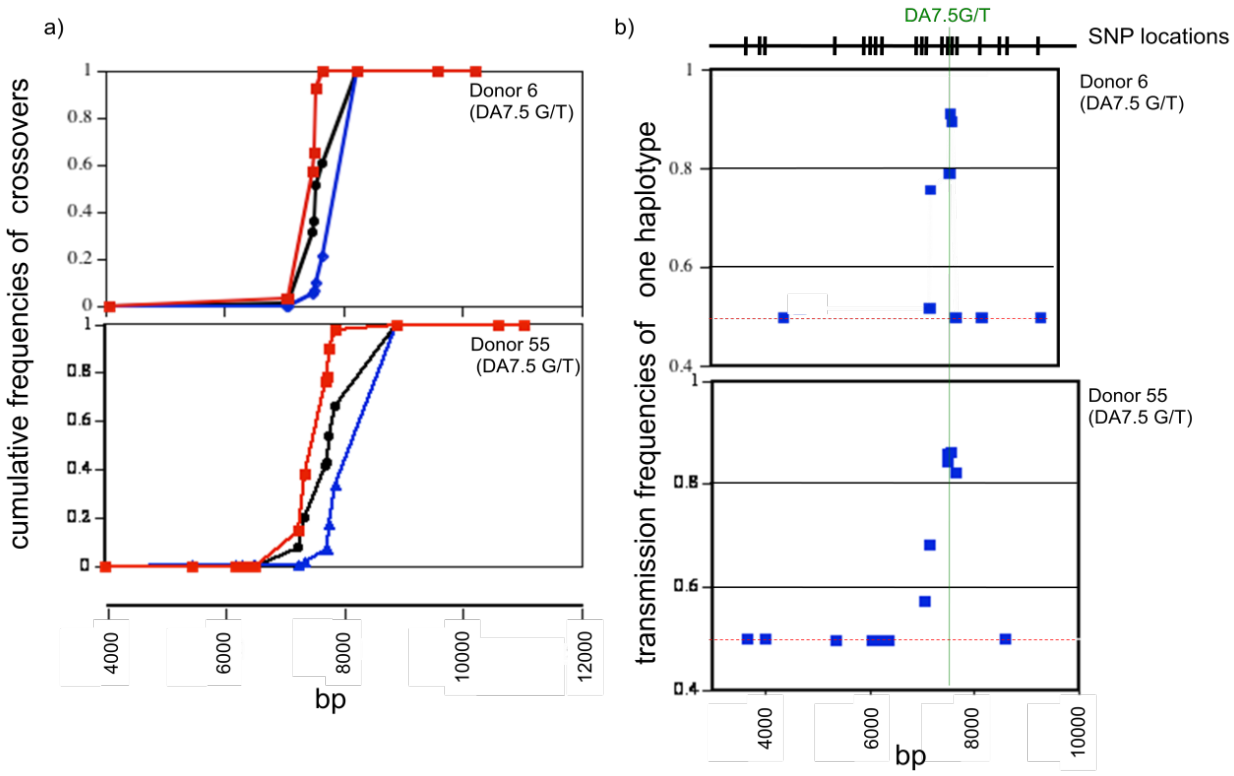

Figure 4. Cumulative frequencies of crossovers (a), and transmission of alleles from one haplotype into crossover progeny (b) in donors 6 and 55. a) The cumulative frequency values for orientation A (red), and B (blue) for donors are shown. The best fit curve for orientations A+B (black) are shown for each donor. The curves for both donors show a similar shape. Cumulative distributions are shifted by 380 bp for donor 6 and 495 bp for donor 55 for A and B crossovers. b) Transmission of alleles from one haplotype into crossover progeny in those donors normalised to equal frequencies of A and B orientations. Massive over-transmission was observed the markers within the hotspot. Central Marker DA7.5G/T showed 80:20 (donor6) and 86:14 (donor 55) over-transmission ratios into crossover progeny (the location of disrupting SNP DA7.5G/T is shown by a green line).

were not located at the centre of hotspot DA. Sperm crossovers in men heterozygous for a motif-disrupting variant show the greatest transmission distortion ratio (~ 90-10) ever seen in a human crossover hotspot (3, 17).

The transmission frequency results showed that the asymmetries in three heterozygous donors (donors 6, 7 and 55) for the motif disrupting SNP DA7.5G/T were very similar in terms of the transmission of DA7.5T to crossover progeny (90:10 transmission ratio) and in terms of the displacement of orientation A versus orientation B crossover distributions (**Fig. 3** and **Fig. 4**). Donor 67 and donor 33 that carried the active G-allele showed the highest recombination frequency respectively of 0.115% and 0.155% among analyzed 6 men. Donors 6 and 55 are both heterozygous for the disrupting SNP and showed similar recombination frequencies (0.07% and 0.06%). However another heterozygous for the disrupting SNP, donor 7 showed markedly different crossover frequencies (0.013%) than donor 6 and donor 55. This gives a clue that there might be other affects (either *cis*- or *trans*- factors) for this donor's hotspot activity.

The direction of biased gene conversion indicates that the chromosome carrying the disrupting allele was suppressed for crossover initiation. This was supported by donor 44, who is homozygous for suppressing allele T for the motif disrupting SNP. There were no single recombinant molecules observed in the crossover assay. The strength of biased conversion in favour of the suppressing allele, combined with the high activity of the hotspot, was sufficient to ensure its eventual population fixation, highlighting the necessarily ephemeral nature of recombination hotspots.

Gene conversion at the Hotspot DA

Gene conversion assays carried out on two men (6 and 55) showed biased non-crossover frequencies to be low, at about 10% of the crossover frequency in both cases (**Table 1**). However to date, observed relative rates of non-cross over (NCO) : cross over (CO) showed major variation between human hotspots (19). However, those studied hotspots showed a wide range of NCO: CO ratios which were typically ~1:3 (19, 20). The non-crossover frequency at hotspot DA seems unusually low, though this might well reflect missing non-crossovers since we only scored those that had co-converted at SNP 7.4 (rs7033234) and 7.5 (rs7036542). There was no significant difference between the two men in the ratio of non-crossovers to crossovers (Fisher exact test, $P = 0.14$). Nevertheless, the data can be used to determine the likelihood that a flanking SNP is co-converted in those non-crossover molecules known to have converted at 7.4 plus 7.5 (**Fig. 4**). While the sample sizes are

Table 1. Poisson-corrected numbers of recombination eventx for donors 6 and 55 are shown above

Donors	Sperm molecules screened	Number of COs	CO frequencies, %	Number of NCOs	NCO frequencies, %	NCOs:COs
6	126,000	192	0.152	15	0.0119	1: 12.8
55	312,000	193	0.062	26	0.0083	1: 7.4

Table 2. Crossover frequencies as estimated for assayed donors

Donor	Population	DA7.5G/T	RF %	lower 95% CI	upper 95% CI
8	British	G	0.196	0.012	0.336
12	French	G	0.203	0.011	0.346
17	British	H	0.032	0.017	0.06
26	British	G	0.361	0.204	0.625
31	British	H	0.045	0.024	0.077
73	British	H	0.02	0.001	0.039
178	Afro-Caribbean	T	0.000	0.000	0.000
180	Afro-Caribbean	T	0.008	0.002	0.019
181	Afro-Caribbean	T	0.000	0.000	0.000
184	Afro-Caribbean	T	0.000	0.000	0.000
185	Afro-Caribbean	T	0.000	0.000	0.000
186	African	T	0.000	0.000	0.000
211	British	H	0.000	0.000	0.000
232	British	H	0.026	0.013	0.046
236	Zimbabwean	H	0.016	0.008	0.043
238	Zimbabwean	H	0.09	0.053	0.146
243	Zimbabwean	T	0.000	0.000	0.000
244	Zimbabwean	T	0.000	0.000	0.000
247	Zimbabwean	H	0.063	0.035	0.104
251	Zimbabwean	H	0.036	0.017	0.07
253	Zimbabwean	T	0.000	0.000	0.000
256	Zimbabwean	T	0.006	0.001	0.017
259	Zimbabwean	T	0.000	0.000	0.000
261	Zimbabwean	T	0.000	0.000	0.000
269	Zimbabwean	H	0.014	0.008	0.022
272	Zimbabwean	T	0.007	0.000	0.010
279	Zimbabwean	T	0.000	0.000	0.000
280	Zimbabwean	G	0.042	0.023	0.071

28 men were assayed for crossovers. The Table shows respectively donors, the populations of the donors to which they belong, the genotype of each man for the motif disrupting SNP7.5G/T, and recombination frequencies with lower and upper 95% confidence intervals (CI), as determined by a Poisson-correction. For example, donor 8 (G/G, homozygous for the active allele of the disrupting SNP DA7.5) showed a higher recombination rate than donor 17 (G/T (H), heterozygous for the active allele for disrupting SNP DA7.5), and moreover, donor 17 showed a higher recombination frequency than donor 178 (T/T, homozygote for the suppressed allele for disrupting SNP DA7.5).

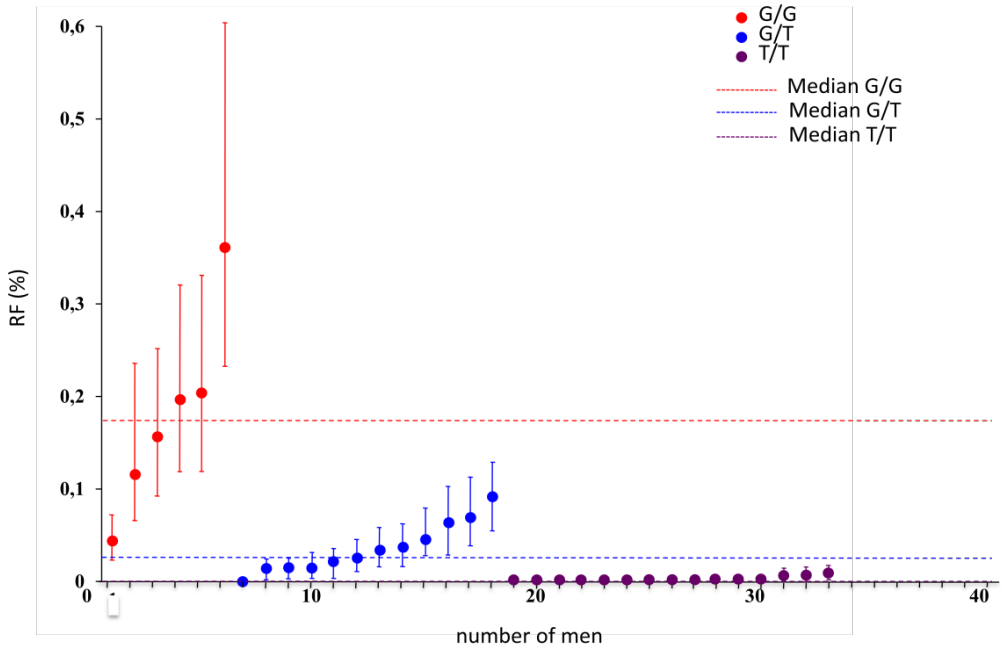

Figure 5. Variation in crossover activity between men at Hotspot DA, indicating their status for the disrupting SNP DA7.5. Individuals were grouped to their status as men homozygous for the active allele (G/G, shown in red) and suppressed allele (T/T, shown in purple), and heterozygous (G/T, shown in blue) for the disrupting SNP DA7.5. 95% Confidence intervals (CI) were estimated using a Poisson correction. Dashed lines indicate the median crossover frequencies observed within each group.

small, there was clear evidence for a very steep gradient, with the likelihood of co-conversion declining rapidly as one move away from the selected sites 7.4 and 7.5. This fits with previous studies of non-crossover distributions in human hotspots.

This suggests perhaps two different pathways for generating NCOs and COs, as seen in yeast and mice. The NCO pathway (maybe the SDSA model (21)) would require that SNP 7.5 is always incorporated into a G/T heteroduplex in G/T heterozygotes and is always repaired in favour of the T-allele during e.g. early mismatch repair. The CO pathway would have to be different, either in terms of a lower likelihood that a G/T heteroduplex is generated, or in terms of a lower visibility of the heteroduplex to mismatch repair or a lower strength of bias in favour of T during repair. Also, a similar story is seen at a PAR2 hotspot (20). Gene conversion that favors the transmission of GC-alleles over AT-alleles, was not observed at Hotspot DA.

To sum up, Hotspot DA is the unique example to date of strongly direct *cis*-regulation for hotspot on/off polymorphism in which a disrupting SNP is located close to its centre. This is confirmed by our reciprocal crossover asymmetry and non-exchange bias gene conversion assays.

Effect of Trans-Regulator PRDM9 on Hotspot DA

To test whether the *trans*-regulator PRDM9 has any influence on crossover frequencies at Hotspot DA, a more efficient crossover assay was used to analyze crossover frequencies. The crossover assay relied on nested repulsion-phase allele-specific PCR to selectively amplify crossovers from sperm DNA. The principle of the crossover assay had been explained in detail in Ergoren, 2013 (18), but in summary, two sets of heterozygous selector sites located within LD blocks flanking both sides of the hotspot are identified and optimized for primers to allow for efficient crossover detection in as many men as possible. The crossover frequency results of multiple men would give clues about the factors that influence hotspot activation in *cis* and in *trans*. To allow as many PRDM9 variants as possible to be tested, 74 Afro-Caribbean and Zimbabwean (African) and 156 European men were investigated (**Table 2**).

Twelve men who were homozygous for the suppressed allele (T/T) of the disrupting SNP DA7.5 and who had different *PRDM9* non-activating variants (C, E, L4, L6, L7, L11, L12, L13, L15, L21, L22) were studied to see whether any of the *PRDM9* variants can rescue the suppression caused by the suppressed T SNP allele acting in *cis* (**Fig. 5**). The *PRDM9* variant L4 generally binds five of the exact 8 bases (10) of the 13-bp motif, but one of its exact bases is the disrupting SNP, so L4 (donor 236) variants only bind to four exact bases of the motif and so rescued only a few crossovers (~1-10 positive crossover reactions in 100,000 molecules). Interesting results come with the men carrying A and L4 variants (donors 180 and 280). These men had different disrupting SNP genotypes (donor 180 (T/T) and donor 280 (G/G)), and it was the men homozygous for the active allele (G/G) of the disrupting SNP who showed higher recombination frequencies. However, d280 and d236 were both homozygous for the suppressed allele (T/T) of the disrupting SNP, and showed very low hotspot activity when compared to the other men who were homozygous for the suppressed allele. Previously, the L20 variant had been observed to be non-activating at Super-hotspots F and U (10). Conversely, L20 has

been shown to activate the hotspot MSTM1b. The PRDM9 variant L20 therefore has entirely different activation profiles between different super-hotspots (10). Despite donor 211 is a heterozygous donor for SNP DA 7.5, Hotspot DA was inactive. However, donor 211 is heterozygous for two PRDM9 alleles (L20 and E) and turns off the activity of Hotspot DA. Even though it is not that strong, perhaps the *PRDM9* L4 variant has an activation effect on Hotspot DA. Recently, a hotspot activated by C variants of *PRDM9* was studied by Berg *et al.,* 2011 (10). The C variant of *PRDM9* totally inactivates Hotspot DA (**Table 3**).

Despite a significant association of *PRDM9* status with crossover frequencies, a large degree of variation can be observed between men with the same *PRDM9* status. Men who carry A/A alleles show variation according to their disrupting SNP genotype, and when considering that the *cis*-effect of

Table 3. Sperm crossover frequencies and their relationship with *PRDM9* status

donor	population	PRDM9 allele 1	PRDM9 allele 2	DA7.5G/T	RF %	lower 95% CI	upper 95% CI
44	British	A	A	T	0.000	0.000	0.000
256	Zimbabwean	A	B	T	0.006	0.001	0.017
272	Zimbabwean	A	B	T	0.007	0.000	0.010
73	British	A	A	H	0.020	0.001	0.039
31	British	A	A	H	0.045	0.024	0.077
247	Zimbabwean	A	A	H	0.063	0.035	0.104
6	British/Indian	A	A	H	0.069	0.004	0.135
238	Zimbabwean	A	A	H	0.090	0.053	0.146
67	British	A	A	G	0.155	0.007	0.302
8	British	A	A	G	0.196	0.012	0.336
12	British	A	B	G	0.203	0.011	0.346
26	British	A	A	G	0.361	0.204	0.625
261	Zimbabwean	L21	A	T	0.000	0.000	0.000
181	Afro-Caribbean	A	L11	T	0.000	0.000	0.000
243	Zimbabwean	B	L15	T	0.000	0.000	0.000
256	Zimbabwean	A	L12	T	0.000	0.000	0.000
244	Zimbabwean	L22	A	T	0.000	0.000	0.000
178	Afro-Caribbean	C	A	T	0.000	0.000	0.000
184	Afro-Caribbean	A	L7	T	0.000	0.000	0.000
180	Afro-Caribbean	L4	A	T	0.008	0.002	0.019
7	British	A	E	H	0.013	0.007	0.025
269	Zimbabwean	A	L11	H	0.014	0.008	0.022
232	British	A	E	H	0.026	0.013	0.046
17	British	A	L20	H	0.032	0.017	0.060
251	Zimbabwean	A	L16	H	0.036	0.017	0.070
280	Zimbabwean	L4	A	G	0.042	0.023	0.071
33	British	A	L2	G	0.115	0.060	0.224
186	African	L6	L13	T	0.000	0.000	0.000
253	Zimbabwean	C	C	T	0.000	0.000	0.000
185	Afro-Caribbean	C	L12	T	0.000	0.000	0.000
279	Zimbabwean	L6	C	T	0.000	0.000	0.000
211	British	L20	E	H	0.000	0.000	0.000
236	Zimbabwean	L4	L22	H	0.016	0.008	0.043

Men were ranked with recombination frequencies (RFs), and 95% confidence intervals were determined by a Poisson-correction. *PRDM9* variants and SNP information for each man for the motif disrupting SNP7.5G/T are also shown above. The *PRDM9* variant allele B acts like allele A (Berg *et al.,* 2010). The other variant alleles of *PRDM9* are indicated in blue.

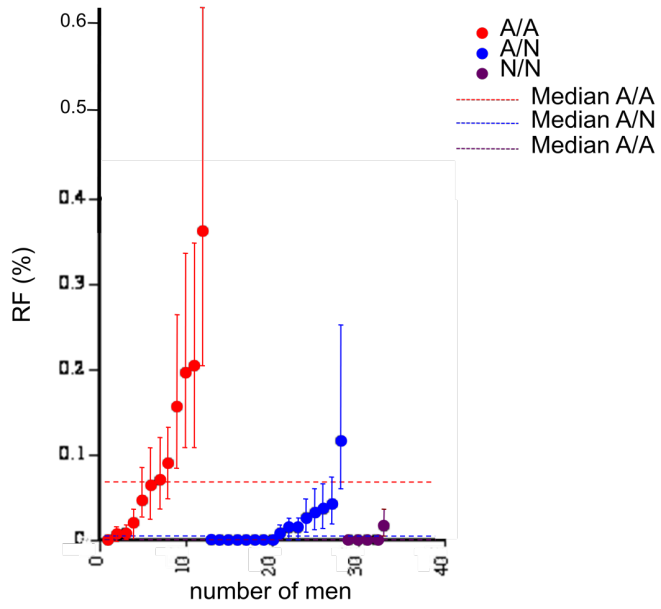

Figure 6. Variation in crossover activity between men at Hotspot DA. Individuals were grouped by their status into men carrying two A-alleles (A/A shown in red), one A-allele (A/N shown in blue) or non-A alleles (N/N shown in purple). 95% Confidence intervals (CI) were estimated using a Poisson-correction. Dashed lines indicate the median crossover frequencies observed within each group.

the variation was as large as 60-fold between donors 26 (G/G) and d256 (T/T). However, the same disrupting SNP genotypes showed a very small, ~3-fold difference between donors d26 (G/G) and donor 67 (G/G), and 4.5-fold between donors 238 (G/T) and donor 73 (G/T) (**Fig. 6**).

However, a comparison of recombination rates between A/A men and A/N men that were heterozygous for the motif-disrupting SNP is significant (one-tailed t-test, P= 0.0153). Thus, the *trans*-regulator *PRDM9* has a significant effect on hotspot initiation in Hotspot DA.

To conclude, both the *trans*-regulator PRDM9 and the *cis*-regulatory disrupting SNP DA7.5 within the 13-bp motif close to the centre of the hotspot, have major influences on hotspot initiation in Hotspot DA. Hotspot DA is the only human crossover hotspot in which the disrupting SNP within the motif has a very strong *cis* effect for the hotspot turn on/off polymorphism, and this *cis*-regulation of the disrupting SNP was confirmed by re-sequencing.

Discussion

About 33,000 hotspots were estimated to be present in the human genome from historical LD block boundaries (5), corresponding to about one in every 50 to 100 kb of sequence. To date, only ~40 hotspots have been directly characterised in individuals by sperm typing (2, 3, 15, 16, 17, 19, 22, 23). In addition to this ~40% of human recombination hotspots have a 13-bp CCNCCNTNNCCNC sequence motif that determines hotspot activity (4).

Myers *et. al.* (5) and colleagues drew up a shortlist of best candidate LD hotspots based on having the 13-bp CCNCCNTNNCCNC motif with a disrupting SNP and good historical activity. This raised the question of whether the nature of the motif offers any clues for understanding the molecular basis of recombination hotspots. Firstly, LD hotspot DA was genotyped using high-resolution genotyping techniques in our European semen donor panel and tested to see if the motif-disrupting SNP lay at the centre of the hotspot as defined by LDU mapping. Confirming the hotspot with LDU mapping is crucial because it confirms the existence of the LD hotspot in our donor panel, the location of the motif with the disrupting SNPs, and it also allows the width of the hotspot to be estimated. LDU mapping confirmed LD hotspot DA to have a width of 1-2 kb which is characteristic of hotspots (1). According to the LDU map, the motif disrupting SNP was at the centre of Hotspot DA, and this hotspot had an estimated historical recombination frequency of ~5×10^{-4}.

Previously, major variation in crossover frequencies had been observed (17), and some of these variations were associated with specific hotspot sequence variants that influence the efficiency of crossover initiation between chromosomes (16, 17, 24). This manifests as reciprocal crossover asymmetry in heterozygotes, as higher initiation on one haplotype can lead to biased gene conversion tracts accompanying the crossover, with markers acquired from the less active homologue (16). In a previously characterised hotspot, both crossovers and non-crossovers indicated comparable biased transmission in both crossover and non-crossover exchange conversion products (17). Therefore, crossovers and non-crossovers were in general influenced by the same factors.

A 15-kb target interval around Hotspot DA that was centered on the motif-disrupting SNP DA7.5 within the 13-bp CCNCCNTNNCCNC hotspot sequence motif was assayed for recombination activity in six donors and the crossover frequency showed Hotspot DA to be a regular hotspot with an average crossover rate (~8×10^{-4}).

Another factor potentially at play is meiotic drive. This is observed in sperm when a recombination suppressing haplotype is over-transmitted to progeny, both at crossovers and at gene conversions, without exchange of flanking markers. Other factors besides the sequence within the hotspot of identical haplotypes could regulate the activity of a presence/absence hotspot polymorphism (17). Therefore, Hotspot DA provides evidence that the motif is likely to be involved in promoting the initiation of recombination. Five assayed donors showed very similar hotspot locations for orientation A and B crossovers. The centre of the hotspot is displaced 3′ to DA7.5 by ~80 bp. Sperm crossovers in men heterozygous for a motif-disrupting variant show the greatest transmission distortion ratio (~90-10, T: G) ever seen in human crossover progeny (24).

In humans (17, 24) different recombination initiation rates result in over-transmission of alleles from the suppressed haplotype into recombinant progeny (24). The highest transmission distortion had been seen in hotspot MSTM1a (on average 72:28 versus 50:50 for Mendelian transmission) for markers

nearest the centre of the hotspot, which is similar in intensity to that seen at the other mammalian hotspots showing distortion (16, 17). Donors that are homozygous for the active G-allele showed the highest recombination rate among the other donors. The direction of biased gene conversion indicates that the chromosome carrying the disrupting allele was suppressed for crossover initiation. This is supported by donor 44, who is homozygous for the disrupting allele (T/T) and shows no recombination (RF = 0%). *Cis*-regulation has previously been seen and has reduced the hotspot activation at studied hotspots (10, 16, 17, 20, 24), but the *cis*-regulation of Hotspot DA has provided the first strong evidence that *cis* factors have a direct influence on the turn on/off activity of hotspots.

Hotspots with high recombination rates also tend to show high non-crossover frequencies. This correlation was seen either between men at a given hotspot or when compared between hotspots (17, 22) at a fairly stable CO: NCO ratio of 2:1. Previously, the CO: NCO ratio was estimated at 4:1 in Europeans and 1:1 in Africans by population genetics approaches (26). In these estimates the influence of gene conversion on LD, which can have detrimental effects in association studies, was ignored. The inflation of historical recombination on LD cannot be explained by crossover-based recombination rates alone (26). Another study by Ardlie *et al*. (2002) was reported non-exchange conversion at a ratio of 3:1-10:1 in favour of non-exchange conversions, with 6:1 being the best estimate (27). This showed incomplete LD that was not readily explained by the expected historical recombination rate based on crossovers alone. It is important to remember that COs are always detected by repulsion-phase PCR, but NCOs can only be detected when they occur in a region including an informative SNP. Therefore, many NCOs are missed, so the quoted ratio may be misleading.

A wide range of CO: NCO ratios have typical values of ~3:1 between men. In this study, the directly observed CO: NCO ratios varied between two analysed donors for Hotspot DA, but at the lower end of what would be expected (10:1). The non-crossover frequency at Hotspot DA seem unusually low, although this might well reflect missing non-crossovers since only those that had co-converted at SNPs 7.4 and 7.5 were scored. There was no significant difference between the two donors in the ratio of non-crossovers to crossovers (Fisher exact test, $P = 0.14$). In contrast, CO: NCO ratios showed a significant and very strong variation between 14 men at hotspot *SPRY3* (20). Hotspot *SPRY3* is located in the minor human pseudoautosomal region (PAR2) and may behave differently to Hotspot DA because either PAR1 or PAR2 must engage in a crossover event at any given meiosis to prevent non-disjunction. Therefore, factors influencing the CO/NCO decision may be functioning differently between pseudo-autosomal and autosomal hotspots. Alternatively, it is equally possible that differences between men at a given hotspot have not yet been detected at autosomal hotspots.

When reciprocal transmission rates of markers in crossovers and non-crossovers were compared for Hotspot DA, the non-crossovers showed complete bias resulting in the transfer of DA7.5T to the DA7.5G haplotype, and with no instances of G>T transfers. This extreme bias is highly significant ($P = 1.2\times10^{-10}$) and is in the same direction as the bias seen in crossovers. For both donors 6 and 55 the strength of the distortion is greater in NCOs than in COs. For each donor, this disparity is not significant (Fisher exact tests on numbers of 7.5 G- and T-containing COs versus numbers of G- and T-containing NCOs). However, it is significant for data from both donors combined, even on raw counts of numbers of NCOs before Poisson correction. It therefore appears that NCOs show a significantly stronger bias towards acquiring 7.5T compared with COs.

Mechanistically, human recombination hotspots can be explained for all observations of recombination events by the single canonical double-strand repair model of recombination (DSBR). The canonical DSBR model has the potential to explain both biases within the same intermediates. In the DSBR model of recombination, both COs and NCOs are produced from the same recombination intermediate and a dHJ by resolution in different planes. Therefore, the intermediates destined to become crossover and non-crossover molecules are perhaps sensed and processed differently by the mismatch repair (MMR) process, resulting in gene conversion with different degrees of bias. This may be manifested by differences in MMR recognising the heteroduplex and/ or differences in the strength of the bias in the lower of T during repair. Alternatively, the extent of resection of the Spo11-induced DSB may differ between molecules destined to become COs and NCOs such that there may be a difference in the likelihood that a G/T heteroduplex is generated at SNP DA7.5.

Evidence for two pathways of non-crossover generation in the human genome comes from the observations of Sarbajna *et al.,* (2012) (20) at the minor pseudoautosomal hotspot SPRY3. They observed at least a proportion of non-exchange conversions being generated via a second, perhaps SDSA pathway (20). This SDSA pathway could be applicable to Hotspot DA. This pathway would require that SNP 7.5 is always incorporated into a G/T heteroduplex in G/T heterozygotes, and is always repaired in favour of the T-allele during, for example, early mismatch repair. Also, meiotic drive in favour of T will lead to hotspot extinction, and not just modest down-regulation as seen at most hotspots.

The identification of PRDM9 as the major *trans*-regulatory factor for specifying and regulating hotspot activity (10, 11), has dramatically increased our understanding as to what controls human crossover hotspots.

Variant *PRDM9* allele A is the most common allele in the European population, and it binds to a 13-bp motif in human LD hotspots identified using European HapMap data (7). Berg *et al.,* (2010) (10). provided evidence for PRDM9 affecting crossover activities in sperm independently of a hotspot motif that had been identified to be the PRDM9 binding site *in vitro* (6). Moreover, additional data addressing PRDM9 regulation and the relationship with the hotspot motif came from

the investigation of recombination hotspots tuned to *PRDM9* C variants, which are more common in Africans (11). However, the LD Hotspot DA was shown to be activated by the *PRDM9* A-variant, and only in men that are homozygous for the active allele (G/G) and heterozygous (G/T) for the disrupting SNP DA 7.5 within the motif. Additionally, the men homozygous for the suppressed allele (T/T) of the disrupting SNP DA7.5 showed either no or extremely low recombination rates in the presence of the *PRDM9* allele A. Men homozygous for the active allele (G/G) of the disrupting SNP and who carried a non-A *PRDM9* variant need to be studied to prove the direct influence of the T allele on the hotspot turn on/off status, but unfortunately, there were no suitable donors in our panel.

The data presented in this study is the only evidence for human crossover hotspot regulation by a very strong *cis*-regulatory disrupting SNP. However, the *trans*-regulator *PRDM9* has a secondary effect. Sperm crossover activity was highly dependent up on *cis*-regulation in Hotspot DA. However, men who carried different *PRDM9* variants showed variable crossover frequencies depending on their disrupting SNP genotypes and *PRDM9* status. In contrast, at some Super-hotspots independent of the presence of the hotspot motif (10), sperm crossover activity was highly dependent on specific *PRDM9* variants. Moreover, a motif-strengthening SNP at Hotspot 5A was associated with a suppressed haplotype (11), and it appears that the presence of a motif is not sufficient for recombination either, as the motif is common and can be seen in recombination cold sequence outside of a hotspot (10). However, a motif-disrupting SNP was found to be associated with the suppressed haplotype at hotspot NID1 (17), which is opposite to Hotspot 5A, where a better-matched motif associated with the suppressed haplotype (4, 5, 11).

Men homozygous for the motif-suppressing allele T at SNP DA7.5 showed no recombination or extremely low recombination frequency. In contrast men homozygous (G/G) and heterozygous (G/T) for SNP DA7.5 showed regular hotspot crossover frequencies depending on their *PRDM9* status. T/T men completely turned off their recombination activity regardless of the presence of the variant *PRMD9* A allele. Therefore, *cis*-regulation is a major factor that controls the hotspot activity in Hotspot DA.

Subtle changes in ZnF arrays between *PRDM9 v*ariants can have a strong effect on recombination activity. The *PRDM9* variant L20 has been shown to activate the hotspot MSTM1b, and show an entirely different activation profile between Super-hotspots (10). Also, the C variant of *PRDM9* showed hotspot activation (11). Additionally, men with variants that have identical binding predictions to the hotspot motif can display variation in crossover frequencies as strong as activation/non-activation (10). Hence, *PRDM9* binding to a specific recognition site would be the most ready explanation for how subtle changes between alleles created non-activating variants incapable of recombination (10). In Hotspot DA, non-activating *PRDM9* variant L4 binds five of the exact 8 bases of the 13-bp motif, including the SNP DA7.5. Therefore, in SNP DA7.5 T/T men, the *PRDM9* L4 variant is predicted to bind to 4 bases for the 13-bp motifs and activate hotspots with extremely low recombination frequency.

Finally, all active and suppressed haplotypes were determined for Hotspot DA, and the disrupting SNP DA7.5 within the 13-bp motif sequence was found to be the only *cis*-regulator and major factor for hotspot activation in this hotspot. Recently, The first extensive analysis of the in vivo binding of PRDM9 in two mouse strains that express PRDM9 variants with different zinc finger arrays is indicated that PRDM9 binds to two classes of sites: sites at recombination hotspots (class 1) and sites that do not showany sign of recombination (class 2) (28). Additionally, Grey *et al.,* (2017) has demostrated that that PRDM9 binding to recombination hotspots (class 1 sites) is Spo11 independent (28). Moreover, active PRDM9 enzimatively is translated to a long-lived complex at specific DNA-binding sites during meiosis prophase I that that might be relevant in stabilizing the components of the recombination machinery to a specific DNA target until DSBs are initiated by Spo11 (29). Thus, our findings prove the importance of the interaction between binding motifs with trans-regulator PRMD9 for the hotspot turn on/off activity.

Final Remarks

Hotspot generation and constant biased gene conversion extinction contribute to the regulation of recombination in the human genome, and this appears to be extremely dynamic. New *PRDM9* variants might create young recombination hotspots, which are then silenced over time through mutation and biased gene conversion. Hotspots would not have to persist given that they are silenced by biased gene conversion, but instead new locations could be activated as recombination hotspots by the generation of new PRDM9 ZnF arrays (10). Despite of the binding specificity and kinetics of the ZnF domain are still not clear. Striender *et al.,* (2017) were carried out two in vitro methods, gel mobility shift assays and switchSENSE, a quantitative biophysical approach that measures binding rates in real time. Therefore, they determined that the PRDM9-ZnF domain forms a highly stable and long-lived complex with its recognition sequence, with a dissociation halftime of many hours (29).

Intensive biased gene conversion, both in to crossovers and non-crossovers, has been found at Hotspot DA. Biased gene conversion that influences crossover and non-crossover activated hotspot activity correlates with *PRDM9* allele A. In Hotspot DA, the lifetime of the hotspot mostly depends on the *cis*-regulatory disrupting SNP DA7.5, and on the *trans*-regulatory factor PRDM9. Recent studies showed that a neutral system of recombination landscape evolution could only be achieved if PRDM9 could evolve at an equal rate or more rapidly than the time it takes for a motif to be completely depleted (30).

Conflict of interest statement

The authors declare there is no conflict of interest.

References

1. Jeffreys AJ, Ritchie A, Neumann R. High resolution analysis of haplotype diversity and meiotic crossover in the human TAP2 recombination hotspot. Hum Mol Genet 2000; 9:725-33
2. Jeffreys AJ, Murray J, Neumann R. High-resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. Mol Cell 1998; 2:267-73
3. Jeffreys AJ, Neumann R. Reciprocal crossover asymmetry and meiotic drive in a human recombination hot spot. Nat Genet 2002; 31:267-71
4. Myers S, Freeman C, Auton A, Donnelly P, McVean G. A common sequence motif associated with recombination hot spots and genome instability in humans. Nat Genet 2008; 40:1124-9
5. Myers S, Bottolo Leonardo, Freeman C, McVean G, Donnelly P. A fine-scale map of recombination rates and hotspots across the human genome Science 2005; 310: 321-324
6. Baudat F, et al. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science 2010; 327:836–840.
7. Myers S, et al. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science 2010 327:876–879.
8. Parvanov ED, Petkov PM, Paigen K. Prdm9 controls activation of mammalian recombination hotspots. Science 2010; 327:835.
9. Hayashi K, Yoshida K, Matsul Y. A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature 2005; 438:374-8
10. Berg IL, Neumann R, Lam KW, Sarbajna S, Odenthal-Hesse L, May CA, Jeffreys AJ PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. Nat Genet 2010; 42:859-63
11. Berg IL, Neumann R, Sarbajna S, Odenthal-Hesse L, Butler NJ, Jeffreys AJ. Variants of the protein PRDM9 differentially regulate a set human meiotic recombination hotspots highly active in African populations. Proc Natl Acad Sci USA 2011; 108:12378-83
12. Maniatis N, Collins A, Xu CF, McCarthy LC, Hewett DR, Tapper W, Ennis S, Ke X, Morton NE. The first linkage disequilibrium (LD) maps: delineation of hot and cold blocks by diplotype analysis. Proc Natl Acad Sci USA 2002; 99:2228–2233
13. McVean GA, Myers SR, Hunt S, Deloukas P, Bentley DR, Donnelly P. The fine-scale structure of recombination rate variation in the human genome. Science 2004 304:581–584.
14. Kauppi L, May CA, Jeffreys AJ. Analysis of meiotic recombination products from human sperm. Methods Mol Biol 2009; 557: 323–355.
15. Webb AJ, Berg IL, Jeffreys A. Sperm cross-over activity in regions of the human genome showing extreme breakdown of marker association. Proc Natl Acad Sci USA 2008; 105: 10471–10476.
16. Jeffreys AJ, Kauppi L, Neumann R. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nat Genet 2001; 29:217–222.
17. Jeffreys AJ, Neumann R. Factors influencing recombination frequency and distribution in a human meiotic crossover hotspot. Hum Mol Genet 2005; 14:2277-87
18. Ergoren MC (2013) Control of meiotic recombination at a human crossover hotspot. University of Leicester, Leicester, United Kingdom.
19. Holloway K, Lawson VE, Jeffreys AJ. Allelic recombination and de novo deletions in sperm in the human beta-globin gene region. Hum Mol Genet 2006; 15:1099-111
20. Sarbajna S, Denniff M, Jeffreys AJ, Neumann R, Soler Artigas M, Veselis A, May CA. A major recombination hotspot in the XqYq preudoautosomal region gives new sights into processing of human gene conversion events. Hum Mol Genet 2012; 21:2029-38
21. Petes TD Meiotic recombination hot spots and cold spots. Nat Rev Genet 2001; 2:360-9
22. May CA, Shone AC, Kalaydjieva L, Sajantila A, Jeffreys AJ. Crossover clustering and rapid decay of linkage disequilibrium in the Xp/Yp pseudoautosomal gene SHOX. Nat Genet 2002 31:272-5.
23. Kauppi L, Jeffreys AJ, Keeney S. Where the crossovers are: recombination distributions in mammals. Nat Rev Genet 2004; 5:413-24
24. Jeffreys AJ, Neumann R. The rise and fall of a human recombination hot spot. Nat Genet 2009; 41:625-9
25. Neumann R, Jeffreys AJ. Polymorphism in the activity of human crossover hotspots independent of local DNA sequence variation. 2006; 15:1401-11.
26. Ptak SE, Voelpel K, Przeworski M. Insights into recombination from patterns of linkage disequilibrium in humans. Genetics 2004; 167:387-97
27. Ardlie KG, Kruglyak L, Seielstad M. Patterns of linkage disequilibrium in the human genome. Nat Rev Genet 2002; 3:299-309
28. Grey C, Clément JA, Buard J, Leblanc B, Gut I, Gut M, Duret L, de Massy B. In vivo binding of PRDM9 reveals interactions with noncanonical genomic sites. Genome Res 2016; 4:580-590.
29. Striedner Y, Schwarz T, Welte T, Futschik A, Rant U, Tiemann-Boege I. The long zinc finger domain of PRDM9 forms a highly stable and long-lived complex with its DNA recognition sequence 2017; 2:155-172.
30. Ponting CP. What are the genomic drivers of the rapid evolution of PRDM9? Trends Genet 2011; 27:165-71