Interdependence of linkage disequilibrium, chromatin architecture and compositional genome organization of mammals

Running title: A common ground for meiotic and mitotic chromatin folding in mammals

Kamel Jabbari*, Johannes Wirtz, Martina Rauscher and Thomas Wiehe

Cologne Biocenter, Institute for Genetics, University of Cologne, Zülpicher Straße 47a, 50674 Köln, Germany.

*Corresponding author: kjabbari@uni-koeln.de; Tel.: +49 221 470 1586

ABSTRACT

Analysing chromatin architecture in interphase nuclei and recombination maps of human and mouse, we observed that blocks of elevated linkage disequilibrium tends to coincide with topologically associated domains (TADs) and isochores. There is a strong correlation between the GC level of TADs, double strand break (DSB) frequency and the local recombination rate. In particular, cold and hot spots of recombination tend to fall in AT- and GC-rich TADs, respectively. Also, binding of proteins which are critical for meiotic recombination hot spots (PRMD9, Spo11, DMC1 and H3K4me3) is positively correlated with the GC level of TADs. We conclude that the occurrence of meiotic DSB and recombination is associated with the same (com)positional features that constrain the architecture of chromosomes in the interphase nucleus of progenitor germ cells or pre-leptotene spermatocytes. This raises the possibility that regional variation of recombination is defined by compositional and epigenetic factors underlying chromatin architecture.

Key words: Recombination frequency, meiosis, chromatin loop, DNA sequence, epigenetic program.

INTRODUCTION

A relationship between genome organization and recombination was discovered from banding of human chromosomes several decades ago. It was shown that translocations are not randomly located on chromosomes and that R bands and G/R borders were the sites of DNA exchanges and included the "hot spots" of mitotic chiasmata (Hecht 1988; Kuhn and Therman 1986). These observations suggested an association with compositional discontinuities (change of GC % along chromosomes) (Bernardi 1989; Ikemura and Wada 1991; Holmquist, 1992), which was later reported (Eyre-Walker 1993a) to be statistically significant. It was proposed that active chromosome domains could be more accessible to the base-mismatch repair system, investigated by Brown and Jiricny (1987, 1988), which preferentially changes mismatched A's and T's into G's and C's. Such mismatch repair, acting during recombination and recognized as gene conversion, would drive recombinogenic domains to GC richness. GC-biased repair of gene conversion (BGC) was suggested as the most likely explanation (Holmquist 1992; Eyre-Walker 1993b). With accumulating data from the human genome project, local rates of recombination were later shown to be positively correlated with regional GC levels in the human genome (Fullerton et *al.* 2001). Regions of weak and strong linkage disequilibrium (LD) were found to be remarkably consistent across human populations suggesting that GC%, DNA polymorphism and repeat content are strongly associated with the local extent of LD, where regions of strong LD are typically GC-poor (Smith et *al.* 2005). BGC is considered by some authors to be the inducer of the correlations between local GC % (isochores) and recombination rate (Duret and Galtier 2009; Weber et *al.* 2014). Other authors argue that local GC-richness may be the driving force for recombination (Marsolier-Kergoat et *al.* 2009). Whole genome sequencing of tetrad products from *Saccharomyces*, *Neurospora*, *chlamydomonas* and *Arabidopsis* did not show GC-bias, implying that a GC% *vs.* recombination correlation is unlikely to be explained by gene conversion (Liu et *al.* 2018).

It was shown (Blat et *al.* 2002) that meiotic chromosomal protein loading is modulated by isochores and that R-bands (GC-rich isochores dense regions) differentially favor double-strand break (DSB) formation during meiosis. In mice and humans, spermatocytes begin to enter leptotene, the first stage of meiotic prophase and later in this stage chromosomes are organized in alternating domains of greater and lower DSB activity (Lichten and de Massy 2011; Grey et *al.* 2011; Grey et *al.* 2017). But features that influence DSB activity at the scale of chromosomal domains, such as chromatin loops (also called contact domains), are still poorly known. It is widely accepted that the binding of transcription factors and chromatin modifiers influences DSB density (Lichten 2008) and that DSBs occur in regions of accessible chromatin that are present in mitotic as well as in meiotic cells. Along with meiotic DSB repair, the search for homology and the catalysis of strand exchange are likely to be spatially and temporally coordinated to allow for the build-up of higher order chromosome structures (for example, synaptonemal complex, chromosome axes and chromatin loops); this is required for the recombination event to successfully join homologues (Zikler and Kleckner 1999; Neal and Keeney 2006; Kleckner 2006). Cohesin has an architectural role in the organization of interphase chromosomes and similar roles have been proposed for cohesin and the related condensin complexes in meiotic and mitotic chromosomes (Tedeschi et *al.* 2013; Ono et *al.* 2003).

The entry of pre-leptotene spermatocytes (PLS) into meiosis is accompanied by several epigenetic changes, in primates as well as in mice. The location of most DSBs is correlated with the trimethylation of histone 3 lysine 4 (H3K4me3) by the DNA binding enzyme PRDM9 (Hayashi et *al.* 2005; Baudat et *al.* 2010; Brick et *al.* 2012). In mouse, 200 DSBs occur per meiosis, of which 15 to 35 will lead to crossovers (COs), and the remainder to non-crossovers (NCO) products (Handel and Schimenti 2010), indicating an excess of DSB events. These additional DSBs serve as facilitators of the process of homology searching and pairing but must be repaired as precisely as those designated to become crossovers (Gray and

Cohen 2016). The outcome is a heterogeneous recombination chromosomal profile where recombination is suppressed in specific mega-base sized regions (Smagulova et *al.*, 2011).

To gain a comprehensive large scale view on the recombination chromatin landscape and dynamic, accumulated large numbers of CO events between the founder haplotypes over successive generations, and high density genotyping or genome-wide sequencing, can be used to obtain accurate localization of these events in human and mouse. The recombination map of accumulated CO events in mice (Margot et *al.* 2017) and LD-blocks from human population are put together to address questions regarding the link between recombination and chromatin architecture, in particular Topologically Associated Domains (TADs) (Dixon et *al.* 2012).

Lately LD-blocks and TADs have also been investigated in two other studies which were deposited on bioRxiv. The first study (Greber et *al.* 2018) did not find a significant correlation of TAD boundaries with classical LD blocks, but when applying a new measure, called ”Linkage Probability”, the authors could report a strong association. The second article (Whalen and Pollard 2018) denies the existence of such an association.

MATERIAL AND METHODS

Recombination and LD-blocks data sets

Multiparent populations such as the Diversity Outbred (DO) mouse stock represents a valuable resource of large numbers of crossover events accumulated in present day sibs from founder haplotypes. It thus allows for an exhaustive exploration of the recombination landscape. We used the data from Morgan et *al.* 2017, a high-density genotype data set from 6886 DO (diversity Outbred) mice spanning 16 breeding generations, in which 2.2 million CO events were localized in intervals with a median size of 28 kb. The construction of linkage maps in pedigrees assumes that every CO is distinct, and can be attributed to one of two specific meioses (in the case of unknown phase). By excluding shared by descent COs, the

authors reduced the total dataset of 2.2 million COs to a set of distinct COs (n=749.560). Coordinates and haplotypes of distinct COs as well as cold spots location are obtained from Morgan et *al.* 2017.

Linkage disequilibrium blocks in human populations are based on the classical r^2 metric used to estimate if a genetic variant is in LD with another genetic variant (Pritchard and Przeworski 2001; Hill and Robertson 1968). LD is defined as the nonrandom association of alleles (of a given gene for example) at distinct close by loci. For our purpose, we used LD-blocks identified with a recently published method (Berisa and Pickrell 2016) applied to sequencing data from European (CEU, TSI, GBR, FIN and IBS), African (YRI, LWK and ASW) and East Asian (CHB, JPT and CHS) populations in the 1000 Genomes Phase 1 dataset (Auton et *al.* 2015). Covariance matrix was separately computed in the European, East Asian, and African meta-populations. The mean block size of 10 000 SNPs is set and used by the algorithm to define the block boundaries. To define human LD-blocks from 1000 genomes dataset, two sets of SNPs are defined as 'approximately independent' if the pair wise r^2 between SNPs in different sets is close to zero. This led to 2605, 1467 and 1725 LD-blocks in African, East Asian and European, respectively. The human recombination map (HapmapII_GRCh37_RecombinationHotspots) was downloaded from the ftp site of the 1000 genome project.

TAD data sets

To study the correlation between chromatin structure and recombination frequency in mammalian cells, we used TAD coordinates from the genome-wide chromatin interaction frequencies (Hi-C experiments) performed on human and mouse embryonic stem cells (Dixon et *al.* 2012). We used the UCSC batch coordinate conversion (liftOver at http://genome.ucsc.edu/cgibin/hgLiftOver) to convert isochore coordinates reported by Costantini et *al.* 2006 from hg18 to hg19 and mouse genome assembly mm9 to mm10 for compatibility with recombination hot/cold spot coordinates. Mouse isochore maps were

visualized with "draw-chromosome-gc.pl" (Paces et *al.* 2004) and Hi-C maps using Juicebox (Durand et *al.* 2016). We also made use of the recently published (Jung *et al.* 2017) Hi-C data from mouse sperm, GEO accession: GSE79230 and mouse embryonic stem cells (Bonev et *al.* 2017), GEO accession: GSE96107.

ChIP-Seq and sequencing data

Meiotic DSBs are induced by dimers of the conserved topoisomerase-like protein Spo11 via a transesterase reaction that links a Spo11 molecule to each 5' end of the broken DNA and release of covalently bound Spo11 to short oligonucleotides (Spo11 oligos). Mapping Spo11 oligos data of mouse spermatocytes are obtained from Lange et *al.* 2016, a total of 13.960 DSB hotspots were defined as regions where the SPO11-oligo is mapping. Genomic coordinates of the strand-exchange protein DMC1-bound single strand DNA from spermatocytes, H3K4me3 and the DNA-binding site of the zinc finger, histone methyltransferase PRDM9 were obtained from Lange et *al.* 2016 and Grey et *al.* 2017.

Statistical analysis

We asked whether the TADs, isochores and LD-blocks overlap more than expected by chance. As a quantitative assessment, we performed an association analysis of genomic regions based on permutation tests using the R/Bioconductor package regioneR (Gel et *al.* 2016). The outcomes of the permutation were subsequently evaluated with the p-values and z-scores of the test. When performing an association analysis, it is possible to detect associations that are not reflective of boundary proximity, although they may be statistically significant. With the "local z-score" function (Gel et *al.* 2016) one can test whether the association between TADs or isochores and LD-blocks is specifically due to the common boundary positions of the analyzed regions. The main function to perform a permutation test

with regioneR is "*permTest*", which takes a region set (RS), a randomization function and an evaluation function as input and returns the object "*permTestResults"* with the computed p-value and z-score which is calculated iteratively with shifted positions in the RS input set. We calculate z-scores for shifts of 500 kb in 5' and the 3' direction from an original position, the focal site. Plotting average z-scores versus shifted positions, one can observe how the value of the z-score changes when moving away from the focal site: a sharp peak at the centre indicates that the association is highly dependent on the specific genomic coordinates while a flat profile indicates regional or diffuse association.

RESULTS

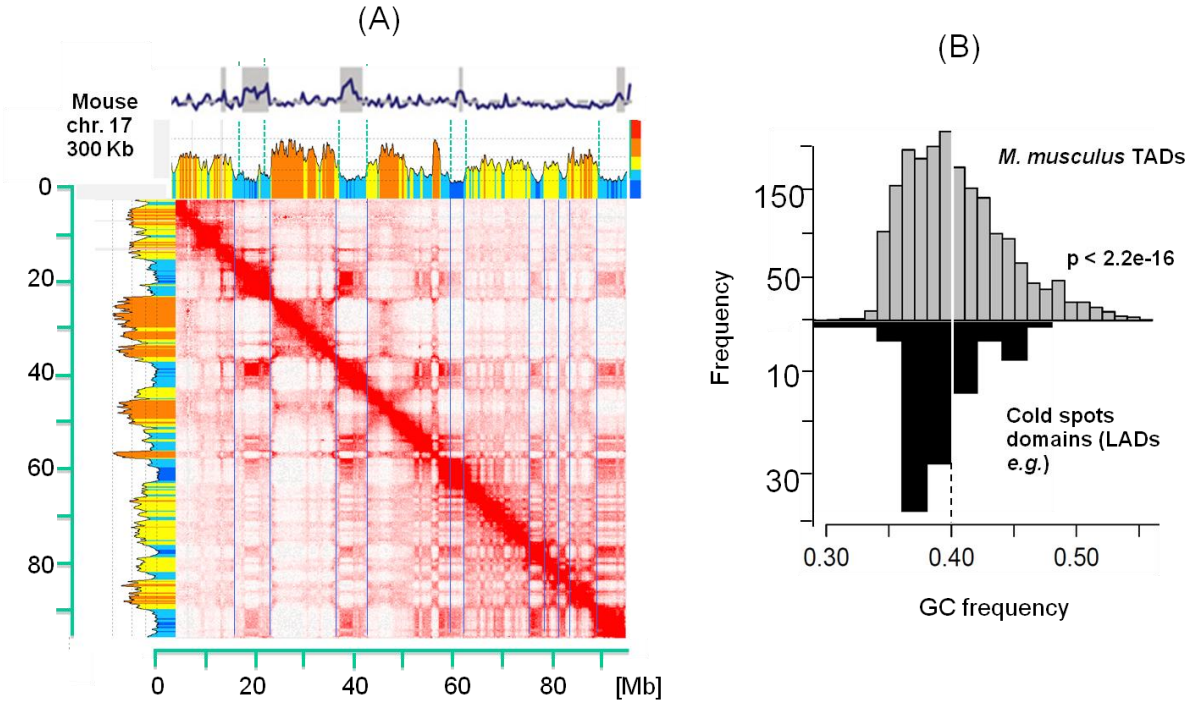
Recombination rate and TADs GC % are strongly correlated

It is known that recombination rates and GC% are positively correlated in the human genome (Fullerton et *al.* 2001), as had been suspected from early chromatin and molecular studies (see Introduction). Here, we revisit this relationship in the light of chromatin TAD organization and density of DSB sites. At the time of DSB formation, most DSBs occur in loop sequences (Blat et *al.* 2002). Hence, the DSB distribution could be affected by loop size (Grey et *al.* 2017; Wang et *al.* 2017) and composition.

To test for a potential relation between TADs and the recombination map, we used the Hi-C map constructed from round spermatids (Jung et *al.* 2017) and a map of recombination cold spots constructed from a DO mice cohort (Morgan et *al.* 2017). As expected (Jung et *al.* 2017), similar results are obtained with embryonic stem (ES) cells (Bonev et *al.* 2017) (Figure S1). As an illustrative example, Figure 1a shows the co-mapping of cold spot domains, isochores and chromatin domains from mouse chromosome 17. Recombination cold spots tend to fall into GC-poor chromosomal domains, which are also frequently found to be attached to the interphase lamina, the "Lamina Associated Domains" (LADs) (Meuleman et

182 *al.* 2013; Kind et *al.* 2015; Jabbari and Bernardi, 2017). This observation is not limited to
183 mouse chromosome 17. A histogram of GC% of cold spots containing TADs from all
184 chromosomes (Figure 1b) shows that the paucity of cold spot in GC-rich TADs is a general
185 feature (p-value < 2.2e-16).

Figure 1

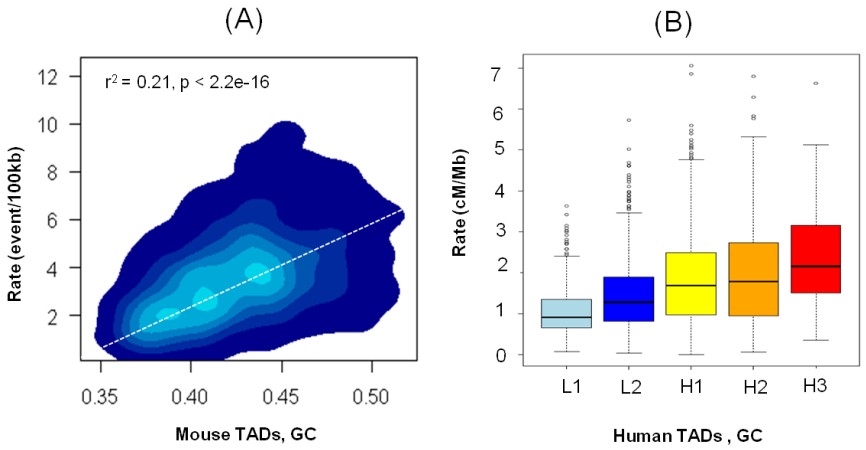


186

187 **Figure 1**. **Spatial correlation between recombination rate, TADs/LADs and isochores**.
188 (A) The heat map of chromatin interactions in mouse chromosome 17 (from Jung *et al.*
189 2017) is aligned to the recombination profile from Morgan *et al.* and to the corresponding
190 compositional profile drawn from mm10 genome assembly, using a sliding window of 300
191 Kb. Increasing GC levels are represented in different colours, deep blue, light blue, yellow,
192 orange and red, respectively; the multi-coloured vertical bars on the top right indicate GC
193 levels that correspond to the compositional boundaries among isochore families. Blue line
194 delimits GC-poor isochores and broken black lines indicate cold spots (in grey), very
195 similar results are obtained with ES cells Hi-C interaction matrices (Figure S1). (B) GC

196 histogram of embryonic stem cells TADs from Dixon *et al.* (top) and cold spots intervals
197 (bottom) defined in Morgan *al* .2017.
198

199 In agreement with the results of figure 1a and 1b, we observe a significant positive correlation
200 between recombination rate and TAD GC% in DO mouse (Figure 2a). The same trend is
201 observed for human, where it is possible to draw a box-plot since isochore boundaries are
202 well defined in this case (Costantini et *al.* 2006; Jabbari and Bernardi, 2017) (Figure 2b).



203

204 **Figure 2**. **Positive correlation between recombination rate and isochore or GC level of**
205 **TADs**. (A) Contour plot representing the correlation between TADs GC level and
206 recombination rate based on data from Morgan et *al.* 2017. (B) Box plot representing
207 TADs GC level correlation with recombination of human using 1000 genome data. L1, L2,
208 H1, H2 and H3 correspond isochore families (for more details see Jabbari and Bernardi
209 2017).
210

211 Judging from the human and mouse analysis, we hypothesize that the correlation between
212 recombination rate, local chromatin architecture and GC% is a general property of
213 mammalian genomes. Binding events of key recombination proteins mapped around DSB

sites are also positively correlated with each other, with recombination rate and with TADs GC % (Table 1).

Table 1: **Correlations between genomic features**. R is recombination event per 100 Kb; blue and black respectively refer to RJ2 and B6 mouse strain data from Grey *et al.* 2017; red refers to data from Lange *et al.* 2016 on B6 mouse strain. All *p*-values are < 2. 2e-16 except for spo11 vs. Prdm9 (r = 0. 16, *p* = 3. 31e-13).

	TADs GC	R	Spo11	Prdm9	H3K4me3	DMC1
TADs GC		***0.46***	**0.38**	**0.25-0.32**	**0.35-0.43**	**0.26-0.20**
R			**0.38**	**0.34-0.39**	**0.22-0.38**	**0.33**
Spo11				**0.86**	**0.24-0.20**	**0.49-0.21**
Prdm9					**0.40-0.55**	**0.21-0.14**
H3K4me3						**0.61-0.41**

Linkage disequilibrium blocks match TADs and LADs

The boundaries of recombination cold spots and TAD domains appear to overlap, as visualized in the coarse-scale heat map shown in Figure 1a. To corroborate this visual evidence quantitatively we tested for boundary overlap between blocks of linkage disequilibrium (LD), TADs, and isochores using data from human 1000 genomes project. The statistical tests for overlap between TADs or isochores and LD-blocks are all significant (Figure S2). There is indeed a clear non-random association (all *p*-values are = 0.001) between the two sets of genomic intervals (isochores or TADs and LD-blocks). Because overlap does not mean boundary proximity, we analyzed the overlap of genomic ranges taking into account boundary matches. To assess the strength of boundary sharing between LD and TAD blocks, as a reference, we first determined the boundary sharing z-score profile for LD blocks of European and African human populations (Figure 3).

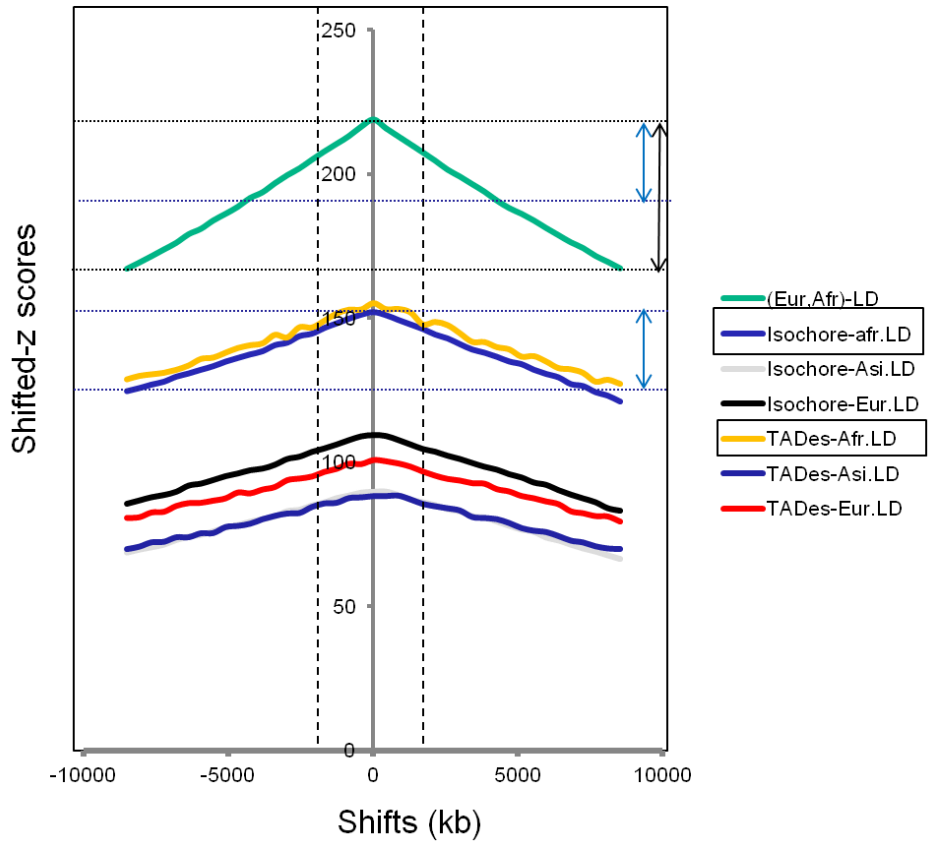


Figure 3. TADs, Isochores and LD-blocks boundaries concordance. Shifted z-score (on the y-axis) changes when moving the RS interval (x-axis): the peak at the centre indicates that the association is dependent on the genomic coordinates, while a flat profile will indicate that the association is regional. Abbreviations: asi. LD, afr. LD and eur. LD correspond to Asian, African and European linkage disequilibrium blocks, respectively. Black and blue arrows show the difference in drop of z-scores between afr.LD *vs*. eur.LD and isochores or TADs *vs*. afr.LD, pointing to a weaker drop of z-scores in the latter compared to the former.

As reflected in z-score profiles (Figure 3), the concordance between TADs or isochores and LD-blocks, although weaker in strength (~50% drop in z-score compared to the reference , see arrows in figure 3), is evidenced by the peaked z-score profiles in all comparisons which suggests that the LD-blocks tend to overlap or match TADs and isochores.

DISCUSSION

Interphase-leptotene chromatin reorganization

Leptotene chromatin domains are seen as linear arrays of chromatin loops, connected by the synaptonemal axis, quite different from the interphase chromatin, where about 40% of the genome is made up of LADs. Can the pre-existing higher order chromatin structure in the interphase of PCG or PLS be related to those of early leptotene? As suggested by Figure 1a**,** the spatial correlation between recombination and loop structure of spermatids and ES cells appears to argue in favour of this link. This is consistent with the lower size of the loops attached to the SC in the telomeric and subtelomeric regions of mouse (Heng *et al.* 1996), known to be GC and gene-rich (Saccone *et al.* 1992), and with the observation that the yeast transcriptional landscape during meiosis dictates the preferred attachment sites of the SC axial elements (Sun *et al.* 2015). A parsimonious explanation could be that the chromatin loops of many leptotene chromosomes are, in part, inherited from the interphase of PGC or PLS through DNA composition constraints and active (*e.g.* H3K4me3, H3K9ac and H3K4me2) and/or repressive (*e.g.* H3K9me3, H3K27me3, H3K9me2) chromatin marks (Tang *et al.* 2016, for a review). The latter could maintain these regions in a "poised" folding state and may happen with the contribution of protein complexes still attached to the LADs DNA released from the nuclear membrane (*e.g.* H3K27me3, H3K9me2), and participating later in the build up of chromosome axis and loops. The majority of recombination cold spots are in constitutive LADs (the GC-poorest isochores), which are enriched for structural variations, in particular deletions (Jabbari and Nürnberg 2016; Morgan *et al.* 2017), that can locally suppress recombination due to local homologous chromosomes miss-alignments disrupting synapsis (Morgan *et al.* 2017). Interestingly, artificial tethering of DSB sites to the nuclear lamina causes a shift from repair by homologous recombination to repair by non-homologous

end-joining (Lemaitre *et al.* 2014). As a consequence, lesions in GC-poor TADs and loops, characterized by lower recombination, may not depend on homologous recombination for their repair. In this context TAD loops will be stabilized as in mitotic chromosomes, by the dynamic binding of cohesin and meiotic insulators. Since BORIS (or CTCF-like), a genomic neighbour of Spo11 (Jabbari *et al.* 2018), is present in male germ cells during and after meiosis, it was proposed that it may interact with at least one of the meiosis-specific subunits of cohesin complexes to contribute to a progressive re-establishment of genome architecture in haploid post-meiotic round spermatids (Lobanenkov *et al.* 2017).

The trade-off between physico-chemical constraints at the DNA level and protein binding partners is expected to affect the conformation of the chromatin fibre that is believed to be primarily determined by its own stiffness, which favours or restricts the formation of long-range contacts (Kleckner 2006). Equally interesting in this regards are the compositional constraints that shape DNA bendability (Vinogradov 2003, 2017) and super-coiling (Naughton *et al.* 2013).

LD-blocks and chromatin neighbourhoods

The relative concordance between TADs or isochores and recombination or LD-blocks revealed in this work has two consequences; (1) it suggests that regional variation of recombination is topologically defined in concert with an underlying compositional and epigenetic framework. GC-rich domains (isochores) are recombinogenic, which may explain the small size of GC-rich LD-blocks and TADs sizes (Figure S3). As a result, conserved LD-blocks, namely those that are shared between Europeans, east Asians and Africans, are significantly larger in size (Figure S4) and their lower recombination rate is in agreement with the strong LD blocks being CG-poor (Smith et *al.* 2005) and enriched in cold spots as shown in figure 1a and b; (2) strong LD between a pair of SNPs in GC-rich recombinogenic TADs may hint to functional chromatin contacts maintained by purifying selection. The ability of LD-blocks to encompass long range SNP interactions in regions with enhanced intra-loop contacts, highlights how allelic chromatin topology analyses can help to infer mechanisms by which SNPs associate with disease and traits (Tang et *al.* 2015).

The Loop organization of the chromatin imposes structural constraints on recombination

It was previously anticipated that meiotic prophase has evolved directly from the latter stages of the mitotic program (Kleckner 2006). This raises the possibility that when the LADs disassemble (likely forming larger contact domains) and chromosomes undergo structural reorganization into linear arrays of chromatin loops, the pre-existing radial loop configuration keeps its “genomic footprints” on the chromatin loop/axis before the recombination process is initiated. Because GC-poor TADs/LD-blocks are longer and gene poor (richer in genes with large introns), one may conclude that the loops emerging from the synaptonemal complex are

of different sizes and may be influenced by chromatin packing within loops, leading to different loop lengths (Zickler and Kleckner 1999).

Nucleosome formation generally restricts the accessibility of proteins to DNA, including Spo11. Meiotic DSBs are reportedly introduced on the chromatin loop regions that transiently interact with the lateral elements of the synaptonemal complex (Blat et *al.* 2002; Aquaviva et *al.* 2013; Panizza et *al.* 2013), suggesting that these chromatin regions may contain nucleosome-depleted regions (GC-rich open chromatin) (Getun et *al.* 2010; Kobayashi et *al.* 2016; Grey et *al.* 2017; Yamada et *al.* 2017). Thus, the local effect is expected to be stronger in GC-rich TADs, where loops are shorter (Jabbari and Bernardi, 2017), nucleosomes are spaced and PRDM9 and CTCF binding sites are enriched (Grey et *al.* 2017).

Evolutionary considerations

As mentioned in the Introduction, whether recombination determines GC level or whether sequence composition drives recombination is still an open question. Because CpG methylation is lower and CpG shortage is stronger in amniotes than in amphibia or fishes (Jabbari et *al.* 1997), some authors speculated (Belle et *al.* 2004) that a GC-bias in mismatch repair increased in the ancestor of amniotes as a consequence of an increase in the level of CpG methylation. On the same basis, other authors (Fryxell and Zuckerkandl, 2000) hypothesized that cytosine deamination and DNA base composition affect each other, generating a positive feedback loop that leads to divergent genetic drifts to high or low GC%; these authors argued that cytosine deamination must be highly dependent on body temperature. The body temperature hypothesis was early proposed (see Bernardi 2007) to explain the formation of the more stable GC-rich isochores of mammals and birds. As far as the correlation between recombination and GC-TADs or isochores is concerned, the GC homogeneity of fish genomes compared to mammals or birds is

compelling, as fish could have evolved the mammalian composiotional pattern had their genome evolved under BGC regimen. Advances in fish and amphibian genetics and genomics will help to understand which factors could have shaped the biased substitution patterns leading to the current mammalian and avian GC landscapes. Recombination COs and BGC frequencies can also be influenced by environmental and physiological conditions, a well-documented case is the heat-sensitivity of the chromatin axis during meiosis (Morgan CH. et *al.* 2017; Lloyd et *al.* 2017). This leaves space for an interplay between selection and recombination in shaping vertebrate genome organization and nuclear architecture. Conservation of the correlation between recombination rates and GC% in human and mouse very likely reflects the evolutionary conservation of mammalian isochores and TADs (Rao et *al.* 2014; Vietri-Rudan 2015 et. *al.* 2015; Jabbari and Bernardi 2017). Therefore, the recombination landscape must be related to the co-regulation of gene expression or insulated interactions in which chromatin folding (loops and TADs) plays a critical role.

CONCLUSIONS

The frequency of cold and hot spots of recombination and its relation to open chromatin was suspected many decades ago, and largely confirmed later (see introduction and Morgan et *al.* 2017). The recent observation of the correspondence between isochores and TADs (Jabbari and Bernardi 2017) and the large number of available genomic maps allowed us to observe that LD-blocks significantly, albeit weakly, match TADs and isochores. Cold/hot spots of recombination are compartmentalized, they correspond to AT/GC rich LADs/TADs. Binding frequencies of key determinants of meiotic recombination hot spots (PRMD9, Spo11, DMC1, H3K4me3) are positively correlated with TADs GC%. This implies that recombination frequency is associated with the same (com)positional features that constrain the distribution of chromatin in the interphase

nucleus. We conclude that the chromatin loop domains in leptotene is inherited in part from the chromatin conformation in interphase. The physical aggregation of cold/hot spots along chromosomes, expectedly led to the finding of the LD-block/isochore/TADs concordance. The recombination landscape in mammals is tied to insulated interactions in which chromatin folding is crucial. Revealing a new aspect of modular recombination underlying alleles co-segregation may open the way for a better understanding of the mosaic architecture of genome regulation and evolution.

372

373 **References**

374 Acquaviva L, Székvölgyi L, Dichtl B, Dichtl BS, de La Roche Saint André C, Nicolas A,
375 Géli V. The COMPASS subunit Spp1 links histone methylation to initiation of meiotic
376 recombination. Science. 2013; 339(6116):215-8.

377 Belle EM, Duret L, Galtier N, Eyre-Walker A. The decline of isochores in mammals: an
378 assessment of the GC content variation along the mammalian phylogeny. J Mol Evol. 2004;
379 58(6):653-60.

380 Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X,
381 Lv X, Hugnot JP, Tanay A, Cavalli G. Multiscale 3D Genome Rewiring during Mouse
382 Neural Development. Cell. 2017; 171(3):557-572. e24.

383 Brown TC, Jiricny J. A specific mismatch repair event protects mammalian cells from
384 loss of 5-methylcytosine. Cell. 1987; 50(6):945-50.

385 Brown TC, Jiricny J. Different base/base mispairs are corrected with different efficiencies
386 and specificities in monkey kidney cells. Cell. 1988; 54(5):705-11.

387 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP,
388 Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. A global
389 reference for human genetic variation. Nature. 2015; 526(7571):68-74.

390 Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B.
391 PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice.
392 Science. 2010; 327(5967):836-40.

393 Berisa T, Pickrell JK. Approximately independent linkage disequilibrium blocks in
394 human populations. Bioinformatics. 2016; 32(2):283-5.

395 Bernardi G. The isochore organization of the human genome. Annu Rev Genet. 1989;
396 23:637-61. Review.

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(6):791-802

Brick K, Smagulova F, Khil P, Camerini-Otero RD, Petukhova GV. Genetic recombination is directed away from functional genomic elements in mice. Nature. 2012; 485(7400):642-5.

Costantini M, Clay O, Auletta F, Bernardi G. An isochore map of human chromosomes. Genome Res. 2006; 16(4):536-41.

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012; 485(7398):376-80.

Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst. 2016; 3(1):99-101.

Duret L, Galtier N. Biased gene conversion and the evolution of mammalian genomic landscapes. Annu Rev Genomics Hum Genet. 2009; 10:285-311.

Eyre-Walker A. (a) Evidence that both G + C rich and G + C poor isochores are replicated early and late in the cell cycle. Nucleic Acids Res. 1992; 20(7):1497-501.

Eyre-Walker A. (b) Recombination and mammalian genome evolution. Proc Biol Sci. 1993; 252(1335):237-43.

Fryxell KJ, Zuckerkandl E. Cytosine deamination plays a primary role in the evolution of mammalian isochores. Mol Biol Evol. 2000; 17(9):1371-83.

Fullerton SM, Bernardo Carvalho A, Clark AG. Local rates of recombination are positively correlated with GC content in the human genome. Mol Biol Evol. 2001; 18(6):1139-42.

Gel B, Díez-Villanueva A, Serra E, Buschbeck M, Peinado MA, Malinverni R. regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. Bioinformatics. 2016; 32(2):289-91.

Getun IV, Wu ZK, Khalil AM, Bois PR. Nucleosome occupancy landscape and dynamics at mouse recombination hotspots. EMBO Rep. 2010; 11(7):555-60.

Gerber S, Fournier D, Hewel C, Horenko I. Imputation of posterior linkage probability relations reveals a significant influence of structural 3D constraints on linkage disequilibrium bioRxiv https://doi. org/10. 1101/255315 [PREPRINT].

Gray S, Cohen PE. Control of Meiotic Crossovers: From Double-Strand Break Formation to Designation. Annu Rev Genet. 2016; 50:175-210.

Grey C, Barthès P, Chauveau-Le Friec G, Langa F, Baudat F, de Massy B. Mouse PRDM9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination. PLoS Biol. 2011; 9(10):e1001176.

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. 2017; 27(4):580-590.

Handel MA, Schimenti JC. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. Nat Rev Genet. 2010; 11(2):124-36.

Hayashi K, Yoshida K, Matsui Y. A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature. 2005 17; 438(7066):374-8.

Hecht F. Enigmatic fragile sites on human chromosomes. Trends Genet. 1988; 4(5):121-2.

Heng HH, Chamberlain JW, Shi XM, Spyropoulos B, Tsui LC, Moens PB. Regulation of meiotic chromatin loop size by chromosomal position. Proc Natl Acad Sci U S A. 1996; 93(7):2795-800.

Hill WG, Robertson A. Linkage disequilibrium in finite populations. Theor Appl Genet. 1968; 38(6):226-3.

Holmquist GP. Chromosome bands, their chromatin flavors, and their functional features. Am J Hum Genet. 1992; 51(1):17-37.

Ikemura T, Wada K. Evident diversity of codon usage patterns of human genes with respect to chromosome banding patterns and chromosome numbers; relation between nucleotide sequence data and cytogenetic data. Nucleic Acids Res. 1991; 19(16):4333-9.

Jabbari K, Cacciò S, Païs de Barros JP, Desgrès J, Bernardi G. Evolutionary changes in CpG and methylation levels in the genome of vertebrates. Gene. 1997; 205(1-2):109-18.

Jabbari K, Bernardi G. An Isochore Framework Underlies Chromatin Architecture. PLoS One. 2017; 12(1):e0168023.

Jabbari K, Heger P, Sharma R, Wiehe T. The Diverging routes of BORIS and CTCF: An interactomic and phylogenomic analysis. Life. 2018; 8(1).

Jabbari K, Nürnberg P. A genomic view on epilepsy and autism candidate genes. Genomics. 2016; 108(1):31-6.

Jung YH, Sauria MEG, Lyu X, Cheema MS, Ausio J, Taylor J, Corces VG. Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. Cell Rep. 2017; 18(6):1366-1382.

Kind J, Pagie L, de Vries SS, Nahidiazar L, Dey SS, et *al.* Genome-wide maps of nuclear lamina interactions in single human cells. Cell. 2015; 163(1):134-47.

Kleckner N. Chiasma formation: chromatin/axis interplay and the role(s) of the synaptonemal complex. Chromosoma. 2006; 115(3):175-94

Kobayashi W, Takaku M, Machida S, Tachiwana H, Maehara K, Ohkawa Y, Kurumizaka H. Chromatin architecture may dictate the target site for DMC1, but not for RAD51, during homologous pairing. Sci Rep. 2016; 6:24228.

Kuhn EM, Therman E. Cytogenetics of Bloom's syndrome. Cancer Genet Cytogenet. 1986; 22(1):1-18.

Lange J, Yamada S, Tischfield SE, Pan J, Kim S, Zhu X, Socci ND, Jasin M, Keeney S. The Landscape of Mouse Meiotic Double-Strand Break Formation, Processing, and Repair. Cell. 2016; 167(3):695-708. e16.

Lemaître C, Grabarz A, Tsouroula K, Andronov L, Furst A, Pankotai T, Heyer V, Rogier M, Attwood KM, Kessler P, Dellaire G, Klaholz B, Reina-San-Martin B, Soutoglou E. Nuclear position dictates DNA repair pathway choice. Genes Dev. 2014; 28(22):2450-63.

Lemaitre C, Zaghloul L, Sagot MF, Gautier C, Arneodo A, Tannier E, Audit B. Analysis of fine-scale mammalian evolutionary breakpoints provides new insight into their relation to genome organisation. BMC Genomics. 2009; 10:335.

Lichten M, de Massy B. The impressionistic landscape of meiotic recombination. Cell. 2011; 147(2):267-70.

Liu H, Huang J, Sun X, Li J, Hu Y, Yu L, Liti G, Tian D, Hurst LD, Yang S. Tetrad analysis in plants and fungi finds large differences in gene conversion rates but no GC bias. Nat Ecol Evol. 2018; 2(1):164-173.

Lloyd A, Morgan C, Franklin C, Bomblies K. Plasticity of Meiotic Recombination Rates in Response to Temperature in Arabidopsis. Genetics. 2018; Genetics. 300588. 2017.

Lobanenkov VV, Zentner GE. Discovering a binary CTCF code with a little help from BORIS. Nucleus. 2018; 9(1):33-41.

Marsolier-Kergoat MC, Yeramian E. GC content and recombination: reassessing the causal effects for the Saccharomyces cerevisiae genome. Genetics. 2009; 183(1):31-8.

Meuleman W, Peric-Hupkes D, Kind J, Beaudry JB, Pagie L, Kellis M, Reinders M, Wessels L, van Steensel B. Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. Genome Res. 2013; 23(2):270-80.

497 Morgan AP, Gatti DM, Najarian ML, Keane TM, Galante RJ, Pack AI, Mott R, Churchill
498 GA, de Villena FP. Structural Variation Shapes the Landscape of Recombination in Mouse.
499 Genetics. 2017; 206(2):603-619.
500 Morgan CH, Zhang H, Bomblies K. 2017 Are the effects of elevated temperature on
501 meiotic recombination and thermotolerance linked via the axis and synaptonemal complex?
502 Phil. Trans. R. Soc. B 372: 20160470.
503 Naughton C, Avlonitis N, Corless S, Prendergast JG, Mati IK, Eijk PP, CockroftSL,
504 Bradley M, Ylstra B, Gilbert N. Transcription forms and remodels supercoilingdomains
505 unfolding large-scale chromatin structures. Nat Struct Mol Biol. 2013; 20(3):387-95.
506 Neale MJ, Keeney S. Clarifying the mechanics of DNA strand exchange in
507 meioticrecombination. Nature. 2006; 442(7099):153-8.
508 Ono T, Losada A, Hirano M, Myers MP, Neuwald AF, Hirano T. Differential
509 contributions of condensin I and condensin II to mitotic chromosome architecture in
510 vertebrate cells. Cell. 2003; 115(1):109-21.
511 Paces J, Zíka R, Paces V, Pavlícek A, Clay O, Bernardi G. Representing GC variation
512 along eukaryotic chromosomes. Gene. 2004; 333:135-41.
513 Panizza S, Mendoza MA, Berlinger M, Huang L, Nicolas A, Shirahige K, Klein F.
514 Spo11-accessory proteins link double-strand break sites to the chromosome axis in early
515 meiotic recombination. Cell. 2011; 146(3):372-83.
516 Pritchard JK, Przeworski M. Linkage disequilibrium in humans: models and data. Am J
517 Hum Genet. 2001; 69(1):1-14.
518 Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn
519 AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at
520 kilobase resolution reveals principles of chromatin looping. Cell. 2014; 18; 159(7):1665-80.

Saccone S, De Sario A, Della Valle G, Bernardi G. The highest gene concentrations in the human genome are in telomeric bands of metaphase chromosomes. Proc Natl Acad Sci U S A. 1992; 89(11):4913-7.

Smagulova F, Gregoretti IV, Brick K, Khil P, Camerini-Otero RD, Petukhova GV. Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. Nature. 2011; 472(7343):375-8.

Smith AV, Thomas DJ, Munro HM, Abecasis GR. Sequence features in regions of weak and strong linkage disequilibrium. Genome Res. 2005; 15(11):1519-34.

Sun X, Huang L, Markowitz TE, Blitzblau HG, Chen D, Klein F, Hochwagen A. Transcription dynamically patterns the meiotic chromosome-axis interface. Elife. 2015; 4.

Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, et *al*. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell. 2015; 163(7):1611-27.

Tang WW, Kobayashi T, Irie N, Dietmann S, Surani MA. Specification and epigenetic programming of the human germ line. Nat Rev Genet. 2016; 17(10):585-600.

Tedeschi A, Wutz G, Huet S, Jaritz M, Wuensche A, et *al*. Wapl is an essential regulator of chromatin structure and chromosome segregation. Nature. 2013; 501(7468):564-8.

Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. Cell Rep. 2015; 10(8):1297-309.

Vinogradov AE, Anatskaya OV. DNA helix: the importance of being AT-rich. Mamm Genome. 2017; 28(9-10):455-464.

Vinogradov AE. DNA helix: the importance of being GC-rich. Nucleic Acids Res. 2003; 31(7):1838-44.

Wang S, Hassold T, Hunt P, White MA, Zickler D, Kleckner N, Zhang L. Inefficient Crossover Maturation Underlies Elevated Aneuploidy in Human Female Meiosis. Cell. 2017; 168(6):977-989. e17.

547 Weber CC, Boussau B, Romiguier J, Jarvis ED, Ellegren H. Evidence for GC-biased
548 gene conversion as a driver of between-lineage differences in avian base composition.
549 Genome Biol. 2014; 15(12):549.

550 Whalen, S Pollard KS. Most regulatory interactions are not in linkage disequilibrium
551 bioRxiv https://doi. org/10. 1101/272245 [PREPRINT].

552 Yamada S, Kim S, Tischfield SE, Jasin M, Lange J, Keeney S. Genomic and chromatin
553 features shaping meiotic double-strand break formation and repair in mice. Cell Cycle.
554 2017; 16(20):1870-1884.

555 Zickler D, Kleckner N. Meiotic chromosomes: integrating structure and function. Annu
556 Rev Genet. 1999; 33:603-754.

557