Defining meiotic recombination landscape in tomato

Chouaref, J.

Publication date
2021

Citation for published version (APA): Chouaref, J. (2021). Defining meiotic recombination landscape in tomato.
General Discussion
The importance of improving knowledge of meiotic recombination

The global population is estimated to reach 9 billion by 2050 [1, 2] emphasizing the need to increase food security. Improvement of production and sustainability in crop production will largely remain based on classical breeding. Meiosis plays a central role in crop production, because it creates new combinations of alleles by the independent assortment of homologous chromosomes [3, 4, 5, 6] and through crossovers (COs), which are reciprocal exchanges between homologous chromosomes [7]. Enhancements of crops could be provided by introduction of desired (e.g. resistance) genes from related wild species in the crop. Due to sequence divergence, recombination between species (homoeologous recombination) remains limited [8, 5, 9]. As a consequence the introgression of resistance traits present in wild relatives into tomato (*S. lycopersicum*), the second most important vegetable crop [10], is hampered: either the introgression does not occur or the introgression drags along undesired genetic sequences, this phenomenon is called linkage-drag [11, 12, 13, 14]. Despite the central role of meiosis in plant breeding, we still lack a full comprehension of the mechanisms that guide the meiotic recombination events in many species.

Meiotic recombination events occur in 1-2kb wide regions in the genome called hotspots [15, 16, 17, 18, 19, 20, 21, 22, 23]. It is thought that the positions where double strand breaks (DSBs) and crossovers take place are determined in part by the chromatin landscape [24, 25, 26, 27, 28, 21]. One of the aims of this thesis was to determine the landscape of meiotic recombination in *S. lycopersicum*. In Chapter 1, we present the current methodologies available to pursue the determination of the meiotic recombination landscape in different crops. In Chapter 2, we address the challenges of investigating the genome-wide localization of meiotic proteins. We adapted and optimized a protocol for chromatin immunoprecipitation combined with high throughput sequencing (ChIP-seq) to work efficiently with low cell input and proteins that bind chromatin only rarely or briefly. Chapter 4 of this thesis shows the first assay for transposase accessible chromatin [29, 30, 31, 32, 33, 34, 35, 36, 37]. In Chapter 3 of this thesis we used CRISPR-Cas9 to generate recq4 and figl1 tomato mutants. We demonstrate the increase of crossover frequency in interspecific hybrids by manipulating genes involved in the resolution of crossover intermediates and observed a potential dual role of FIGL1 in *S. lycopersicum*.

The two aspects of this thesis complement each other. Mutants are not available in many economically important crops because they have a long generation time or are resistant to transformation. Hence by combining both aspects, this thesis provides a good insight of the meiotic recombination landscape in *S. lycopersicum* and the potential methods to exploit the meiotic recombination landscape.
The meiotic recombination landscape of Solanum lycopersicum

The dynamics of meiosis involves both chromosome conformation and gene transcription in order to establish the necessary CO(s) that insure proper segregation of the chromosomes. Meiotic recombination events remain remarkably tightly controlled in terms of occurrence and location even when the chromosomes are subjected to massive rearrangements [38]. The large number of factors constraining the process, demonstrates the complexity of this phenomenon. SPO11 generates DSBs only at already accessible chromatin loci. The pair PRDM9 and HELLS directs the action of SPO11 [39] in mammals towards specific genomic location to introduce DSBs. Indeed, the chromatin remodeler HELLS forces the histone trimethylase PRDM9 to avoid genic regions, by locally increasing chromatin accessibility outside of genes. PRDM9 subsequently trimethylates these loci, making them targets of SPO11 [24, 40, 41]. This example demonstrates the importance of chromatin accessibility for the regulation of meiotic recombination events.

In the chapter 4 of this thesis, we investigated the chromatin accessibility landscape of tomato prophase I meiocytes. To this end we purified meiotic nuclei using fluorescence activated nuclei sorting (FANS) [42]. FANS enabled us to specifically isolate meiotic nuclei and to discard the organellar DNA, which usually impairs ATAC-seq [43, 44, 45]. We generated high quality chromatin accessibility datasets with high genomic coverage and high reproducibility between the biological replicates, and found that most accessible chromatin regions are located around transcription start sites. By comparing the accessible chromatin regions (ACRs) between meiocytes and somatic cells we observed that half of ACRs were ubiquitous. Interestingly, an equal portion of the ACRs was specific to one or the other tissue, suggesting significant tissue-specificity of chromatin accessibility. However, the tissue-specific ACRs around TSSs were not overrepresented among the tissue-specific expressed genes (unpublished data). This suggests that despite chromatin accessibility, expression of the gene requires other factors to proceed. Yet, meiotic specific ACRs are potential targets of SPO11.

Interestingly, we observed a global increase of chromatin accessibility around certain classes of transposable elements (TEs). This is not surprising, since TE expression is upregulated during meiosis [46, 47, 48, 49, 50]. The observed differences in chromatin accessibility at the repetitive elements might be an unexpected side effect of the reorganization of the chromatin structure during early meiotic prophase, as various steps of compaction and decompaction of the chromatin may facilitate accessibility of certain transposable elements [50]. However, all transcriptomic data presented in this study relied on publicly available datasets of entire flower buds. Several studies showed the complex picture of the meiotic transcriptome [47, 48] and hundreds of differentially expressed genes between anthers and flower buds. Additional RNA-seq analysis may provide supporting evidence for our observations of chromatin accessibility at meiotic specific ACR and TE. The meiosis specific ACRs might originate from the activity of pioneer transcription factors (TFs) or chromatin remodelers [47, 48, 51, 52]. Footprint analysis allows the Identification of chromatin accessibility induced by TFs or chromatin remodelers [53, 54]. As an example, using the algorithm TOBIAS, footprinting analysis revealed the importance of Dux, a transcription factor, for zygotic genome activation in mouse [54]. By using on ATAC-seq data, they could identify the binding of Dux at transposable elements, a necessary event of the zygotic genomic activation in mice and humans [54].
Identification of footprints heavily relies on ChIP-seq information to determine the binding sites of certain epigenetic factors to the certain genomic location [54]. Hence gathering ChIP-seq information is of high importance to decipher the chromatin dynamics during meiosis.

In the chapter 2 we established a ChIP-seq protocol to examine chromatin binding sites of meiosis-specific proteins in tomato. Classical chromatin immunoprecipitation is a powerful tool for investigating the genome-wide distribution of DNA-binding proteins [55]. To apply ChIP to tomato meiocytes, we had to overcome several challenges intrinsic to meiosis. First, the collection of anthers containing meiocytes of the stages of interest relies on manual dissection of the flower buds. The collection of sufficient amounts of the appropriate tissue(s) is extremely time consuming. To collect pachytene anthers for the CO-linked MLH1 ChIP-seq and of leptotene/zygotene anthers for the DSBs DMC1 ChIP-seq took more than 2 years of harvesting from forty tomato plants twice a week. The second difficulty is that the binding of meiotic proteins to chromatin lasts only during particular stages of prophase I, making it important to harvest meiocytes that are in a particular meiotic stage. And final difficulty, meiotic recombination events are rare. From the 250 DSBs events less than 10% ends as COs [56, 17, 57]. For all these reasons, we needed to optimize the ChIP procedure to reduce the huge loss of DNA material inherent to classic ChIP protocols. Chapter 2 presents a hand-out method to perform a successful ChIP on tomato anthers. We describe a detailed procedure to identify the meiotic stage of each individual anther and subsequently performed ChIP with low quantities of input material. The isolated DNA can be used for quantitative PCR (qPCR) or to generate libraries for high throughput sequencing. Unfortunately, we failed to meet the challenge of identifying the sequences linked to MLH1 in tomato pachytene. Yet, our results showed that MLH1 avoids the peri-centromeric regions and binds to the chromatin near the telomeres. However, the biological replicates we obtained from this ChIP-seq were not consistent enough to draw solid conclusions and we preferred to not include them in this thesis. The MLH1 antibody is most likely the reason why the ChIP was not reproducible. Hence, we produced a new tomato-specific antibody but have not yet collected sufficient amounts of pachytene chromatin to repeat the ChIP-seq. The same applied for the DMC1 ChIP-seq that we performed in zygotene anthers. To this day despite the major role of MLH1 and DMC1, only a few ChIP studies in mouse identified DMC1 binding sites [58]. Immunocytological analyses of pachytene chromosomes shows that MLH1 interacts with the chromatin in tomato at only 17.58 loci on average per cell and might remain recalcitrant for ChIP methods because of the rare binding events [59]. In the case of a successful MLH1 ChIP-seq in tomato, it would be necessary to perform as well a MLH3 ChIP-seq in order to trustfully identify the Class I COs since only the couple MLH1-MLH3 marks Class I COs [60, 8]. Using tagged proteins is a valuable alternative the generation of ChIP-grade antibodies for each meiotic proteins. For example, tagged Rec8-cohesin ChIP-seq in A. thaliana revealed anti-correlation of Rec8 chromatin binding sites with DSBs and enrichment of Rec8 in pericentromeric heterochromatin [61]. Notably, this was done in Arabidopsis, in which anther size is even more challenging for ChIP analysis.

Alternative chromatin profiling methods exist and it would be beneficial to use them in the future for the investigation of the meiotic recombination landscape in tomato. The first approach to investigate both DNA-protein interactions and chromatin accessibility may be the utilization of the DAM-ID [62]. It relies on the utilization of the DNA adenine methyltransferase (DAM). An expressed recombinant protein consisting of the protein of interest and the DAM protein can methylate the adenine bases in GATC motifs near the sites of protein-DNA
interactions. These methylated sites can be conveniently cut by the methyl-sensitive restriction endonuclease DpnI. The detection of this methylated adenine by high-throughput sequencing helps the identification of binding sites of the protein of interest. The key advantages of the DAM-ID system over ChIP are that it does not require high quality antibodies for meiosis-specific proteins. In a slight variation of this approach, expression of native DAM protein alone can be used to efficiently map of accessible chromatin regions [63]. However, the Dam-construct needs to be expressed at low-level. The Cleavage Under Targets and Release Using Nuclease (CUT & RUN) is a promising method for any chromatin profiling strategy [64]. It is based on the immobilization of unfixed nuclei on magnetic beads. The protein of interest is subsequently targeted by a specific antibody coupled with a protein A micrococcal nuclease. Upon addition of calcium, the DNA-protein-antibody complex is cleaved by the micrococcal nuclease and diffuses out from the nuclei and can be easily extracted after centrifugation of the nuclei extracted from the supernatant for downstream analysis. This method is robust and provides high-resolution profiles for the binding of transcription factors. Moreover, it reduces the amount of input material that is required, which allows genome-wide analysis of protein binding sites in cells that are difficult to obtain in large quantities, such as stage-specific meiocytes. CUT & RUN requires a relatively low-number of cells compared to ChIP-seq and provides a very low background signal. In contrast to ChIP-seq wherein the whole genome is sonicated and only a fraction is pulled down by the antibody, CUT & RUN cleaves only the DNA fragment associated with the protein of interest. When studying proteins, which bind at low frequency, this improves the signal and the downstream analysis of the experiment.

Through Chapter 2 and Chapter 4 our work contributes to lift the veil on the regulation of meiotic recombination events by offering tools to isolate and study stage specific meiocytes and by identifying the ACR of prophase I meiocytes.
Unleashing meiotic recombination in plants increasing the class II CO pathway

Besides understanding the chromatin landscape, which affects positions where recombination occurs, the interest of many meiosis research labs lies in manipulating meiotic recombination frequency. The number of recombination events per generation remains low in many economically relevant crops [10]. This phenomenon impairs gene shuffling, which is the main pathway to improve crops. Thus, increasing the frequency of COs is of great interest for breeders [4, 5, 65]. The last decade has witnessed many breakthroughs in our comprehension of the regulation of crossing over. More than 80 genes involved in meiosis were discovered [8]. Among those genes, negative regulators of CO frequency were identified [31, 33, 35, 34]. Mutations of recq4 and figl1 unleashed crossovers frequency in Arabidopsis to an increase of 7.8-fold [35]. In Chapter 3 of this thesis, we applied CRISPR-Cas9 in an interspecific tomato hybrid (S. lycopersicum x S. pimpinellifolium) to knockout RECQ4. Mutation of recq4 leads to an increase of CO frequency, which is consistent with a recent study with intraspecific hybrids[36]. We noticed however that in our interspecific tomato hybrids, the mutation of recq4 induced a lower increase of COs than in the intraspecific context of Arabidopsis. The difference might be due to the interspecific background since sequence divergence reduces the effect of the recq4 mutation [66, 36]. Our results show that mutation of recq4 also increases CO frequency in an interspecific hybrid background. Intriguingly our CRISPR-Cas9 mutation of figl1 reduced the CO frequency in the same interspecific hybrid background. This contradicts previous studies in Arabidopsis [33, 35] where figl1 increased CO frequency. We explained this difference by analyzing the structure of FIGL1 and the mutant alleles used in this and other studies. Our CRISPR mutation induced a loss of all conserved domains, whereas the Arabidopsis mutant lacked the AAA-ATPase and VSP4 domains but maintained the FRBD1 domain. We potentially highlighted the dual function of FIGL1. While the AAA-ATPase domain might be involved in an anti-CO function, the FRBD1 domain has been shown to interact with RAD51 [66] and might be sufficient to prevent the formation of aberrant recombinant intermediates that cannot be processed to COs. It would be interesting to induce the same mutation in Arabidopsis and observe whether or not the same phenotype appears. On another point, it would be interesting to measure the synergetic effect of recq4 and figl1 mutations in our interspecific hybrid. Previous studies have shown that the combination of recq4 and figl1 enhances crossover rates [33, 66] Since recq4 and figl1 have opposite effects on CO frequency in our interspecific hybrid, it is difficult to predict the outcome of a double mutation. This approach is unlikely to lead to successful increase of the CO frequency in tomato: we have indeed already noticed some aberrant chromosome structure in the chapter 3 of this thesis upon recq4 mutation and figl1 was reported to induce sterility in tomato and other species [67, 36]. It is important to mention that we did not perform any fertility test.

Focusing on components of the class II CO pathway such as the anti-CO proteins FIGL1 and RECQ4 appears as a good strategy to increase CO frequency and to alter their distribution, since COs produced by the class II CO pathways are exempted of interference. It is also possible to unleash the class I CO pathway. Overexpression of HEI10 leads to a dosage-dependent increase of the crossover frequency in subtelomeric euchromatin of Arabidopsis thaliana. Despite the challenges it would be interesting to generate tomato lines overexpressing meiotic proteins involved in Class I CO such as HEI10, MLH1 and MLH3.
Since its discovery, the CRISPR-Cas9 system has quickly become the most efficient tool for precise genome editing. Despite some challenges this method works well in tomato as described in this thesis. It is however necessary to mention that in Europe plants modified by CRISPR-Cas9 are considered genetically modified organisms and are therefore subject to stringent regulation. Hence due to legal uncertainty, virus-induced gene silencing (VIGS) is an alternative important to mention. VIGS uses the plant defense mechanism to induce knocking down of genes [68]. By using VIGS, one could transiently reduce the expression of anti-CO genes in a very precise manner. To conclude, our works suggest that mutation of recq4 increases COs frequency in an interspecific hybrid. Our observation of the unexpected results for figl1 mutation highlight the complexity of the meiotic regulation. Nevertheless we expect the meiosis field will increase its knowledge in the class II CO pathway.

**Conclusion**

This thesis is a modest contribution to the general knowledge of meiosis in plants. With the knowledge drawn from this thesis, it is possible to perform ChIP-seq on stage-specific tomato meiocytes. Application of this protocol could help to draw the map of the histone modification deposition over tomato meiocyte over the various stages of prophase I. Integrating the histone modification information with chromatin accessibility will help us to determine both potential hotspots and coldspots of meiotic recombination of tomato.

**References**


