Development of single-cell analysis methodologies to investigate segregation and dynamics of defined genomic regions during meiosis and interphase

Dissertation

zur Erlangung des

Doktorgrades der Agrarwissenschaften (Dr. agr.)

der

Naturwissenschaftlichen Fakultät III

Agrar- und Ernährungswissenschaften,

Geowissenschaften und Informatik

Martin-Luther-Universität Halle-Wittenberg

vorgelegt von

Steven Dreissig

Geboren am 15.11.1989 in Wolfen

1. Introduction

1.1. The cell cycle and meiosis: Divergent chromosome structure and function

1.1.1. The cell cycle

The concept of the eukaryotic cell cycle was first described in 1850 where it was shown that cells originate by division of a pre-existing cell (Mayr, 1982). Since then, our understanding of the mechanisms which regulate cell cycle progression and cell division has greatly increased. Furthermore, progress in our understanding of the molecular mechanisms regulating these processes revealed that the principal mechanisms of the cell cycle are conserved among all eukaryotes (Dewitte & Murray, 2003). The mitotic cell cycle encompasses four sequential ordered phases which temporally separate DNA replication and subsequent segregation into two daughter cells. Replication, transcription, and repair of the DNA (S phase) and segregation of the chromosomes into the two daughter cells (M phase, mitosis) are interceded by gap phases (G1 and G2 phase) (Dewitte & Murray, 2003). During these phases, also termed interphase, chromatin is hundred or thousand fold less compacted than during mitosis where chromosomes are highly compacted (Llères *et al*, 2009). It is therefore facilitated that the transcription machinery can access genes, which supports the observation that most transcription occurs during interphase (Schubert & Weisshart, 2015). To ensure proper cell division, it is also important for cells to control that DNA replication was completed before entering mitotic division. In this context, it was found that transitional stages (e.g. from G1 to S phase and G2 to M phase) have emerged as important check points where the majority of regulatory proteins operate (Van't Hof, 1985).

1.1.2. Chromatin dynamics during interphase in plants

Interphase chromosomes are less compacted than mitotic and meiotic chromosomes. However, a certain level of organisation is still maintained. In many eukaryotes, interphase chromosomes are organised in distinct chromosome territories associated with several features regulating transcription, silencing, DNA replication, and a distinct higher-order chromatin structure (Cremer & Cremer, 2001, 2010). Although this is conserved in many species, there are nevertheless some variations in interphase chromosome organisation.

It was proposed by C. Rabl in 1885 that interphase chromosomes are organised in a way that reflects their orientation during anaphase which results in centromeres and telomeres being located at opposite poles (Rabl, 1885). In plants, Rabl-configuration was later confirmed in monocotyledonous plants such as *Triticum aestivum*, *Hordeum vulgare*, and *Avena sativa* as well as in dicotyledonous plants like *Pisum sativum* and *Vicia faba* (Dong & Jiang, 1998; Rawlins & Shaw, 1990). Interestingly, there were others where no Rabl-configuration was found, e.g. *Zea mays*, *Sorghum bicolor*, and *Arabidopsis thaliana* (Dong & Jiang, 1998; Armstrong *et al*, 2001). It was later reasoned that Rabl-configuration is not a result of genome organisation nor large genome size, because it was found present in small (*S. pombe*, *S. cereviseae*) and absent in large (*Mus musculus*) genomes as well as present only in specific rice (*Oryza sativa*) cell types (Funabiki *et al*, 1993; Gilson *et al*, 1993; Manuelidis, 1984; Schubert & Shaw, 2011; Prieto *et al*, 2004; Dong & Jiang, 1998). These examples represent some of the interphase chromatin dynamics differences between different species and cell types. However, it appears there is yet another layer of interphase chromatin dynamics in plants which was observed as varying patterns of spatial association between chromosome arms (Schubert *et al*, 2012). Although chromosome territories seem to be separated from each other in a distinct way, chromosome arms within one chromosome territory appear to be more dynamic. These chromosome dynamics are thought to be associated with transcription, silencing, replication and DNA repair (Schubert & Shaw, 2011; Schubert *et al*, 2012; Tiang *et al*, 2012). For example, heterochromatic regions which are enriched in repetitive DNA are often transcriptionally silenced, wherefore adjacent genes are also silenced due to their transcriptionally silent neighbourhood (Cryderman *et al*, 1999; Gottschling *et al*, 1990; Nimmo *et al*, 1994; Cremer & Cremer, 2001; Fischer *et al*, 2006).

Unraveling the spatio-temporal organisation of the genome in the nucleus is important to understand how genes and non-coding DNA sequences are regulated during plant development and environmental changes. For example, it was shown that prolonged heat stress leads to decompaction of heterochromatin and transcriptional activation of repetitive elements in *A. thaliana* (Pecinka *et al*, 2010). These changes in the structure of heterochromatin might also have an effect on adjacent genes when plants are exposed to heat stress. Furthermore, it was shown that the *Arabidopsis CHLOROPHYLL A/B-*

4

BINDING PROTEIN (*CAB*) locus moves from the interior of the nucleus to its periphery upon its transcriptional activation in response to light (Feng *et al*, 2014a). Therefore, elucidating the spatio-temporal dynamics of defined genomic regions during interphase will be important to understand how plants regulate developmental processes and stress response at the chromatin level, and will have significant implications for plant growth in changing environments.
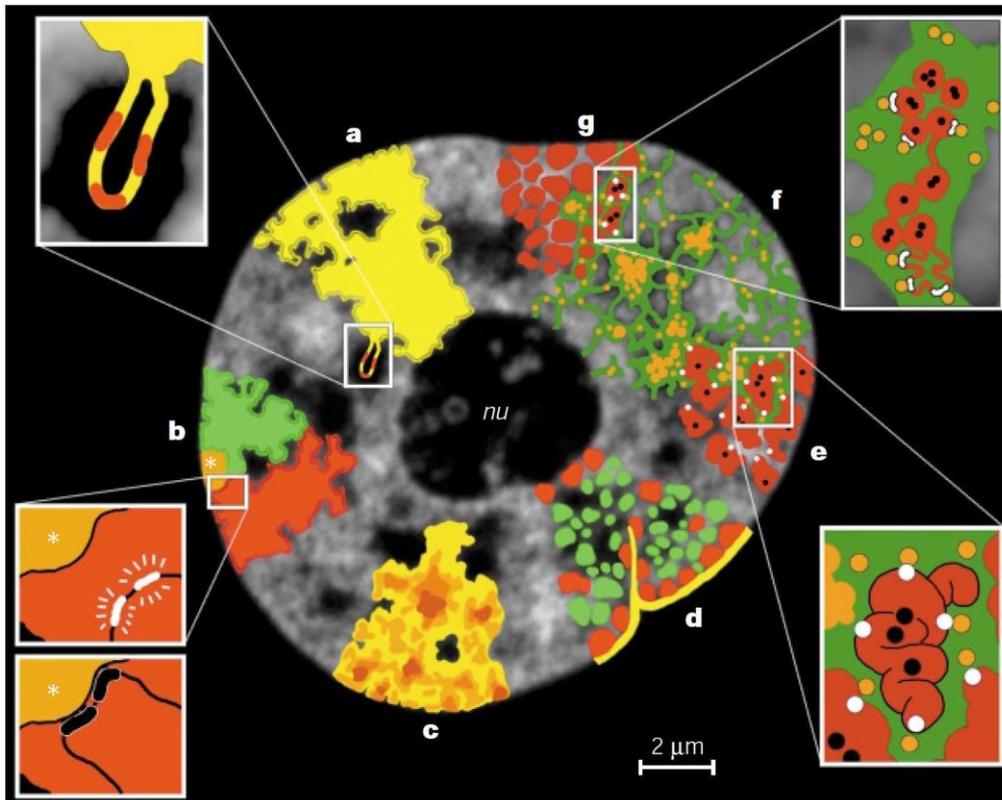

Figure 1 Model of functional nuclear architecture (adopted from Cremer & Cremer 2001). This figure shows structural features of the chromosome-territory-interchromatin-compartment (CT-IC) model. Features are drawn on an optical section of a nucleus of a living HeLa cell. **(a)** CT showing complex folding. Inset shows topological model of gene regulation where a chromatin loop with active genes expands from the CT into an interchromatin compartment. **(b)** CTs contain separate chromosome arm domains (red versus green) and a centromeric domain (orange, asterisks). Inset shows transcriptionally active genes separated from the centromeric region versus silenced genes associated with the centromere. **(c)** CTs contain chromatin of variable density (dark brown, high density; light yellow, low density). **(d)** CT showing early-replicating chromatin domains (green) and mid-to-late-replicating chromatin domains (red) **(e)** Higher-order chromatin structure. Inset shows topological view of gene regulation with active genes (white dots) at the surface of convoluted chromatin fibres and silenced genes (black dots) located towards the interior. **(f)** Interchromatin compartment (green) contains complexes (orange dots) and larger non-chromatin domains (aggregations of orange dots) for transcription,

splicing, DNA replication and repair. **(g)** CT with ~1-Mb chromatin domains (red) and interchromatin compartment expanding in between. Inset shows topological relationships between the interchromatin compartment, active, and silenced genes. The finest branches of the interchromatin compartment end between ~100-kb chromatin domains.

1.1.3. Meiosis and genetic recombination

The majority of eukaryotes reproduce via meiosis, where a diploid cell replicates DNA once and segregates chromosomes twice to generate haploid gametes (Villeneuve & Hillers, 2001). Additionally, during the first meiotic division homologous chromosomes initiate recombination, which can result in reciprocal crossover. Due to the effect of meiotic recombination on genetic diversity it is fundamental to the evolution of eukaryotic genomes and serves as a major tool for crop improvement through breeding. Interestingly, the relative frequency of meiotic recombination is highly variable along chromosomes, with concentration in narrow hotspots (Choi & Henderson, 2015; Choi *et al*, 2013; Saintenac *et al*, 2011). In addition, large regions, for example surrounding the centromeres, are suppressed for meiotic recombination (Salomé *et al*, 2012; Aliyeva-Schnorr *et al*, 2015). In barley (*Hordeum vulgare* L.), interstitial and centromere-proximal regions which are suppressed for recombination contain 12 - 24% of the barley gene complement (Baker *et al*, 2014). Although genetic diversity is reduced in low-recombining regions, they nevertheless contain genes and thus represent a resource that is hardly accessible to plant breeders (Baker *et al*, 2014).

Recombination frequency is controlled at many levels and is initiated by the formation of programmed DNA double-strand breaks (Bergerat *et al*, 1997; Keeney *et al*, 1997). Repair of these double-strand breaks (DSBs) results in either crossovers (CO) or non-crossovers (NCO). COs are reciprocal exchanges of large chromosome fragments between homologous chromosomes whereas NCOs are copies of a small part of the intact chromosome to the broken chromosome without affecting the template chromosome (Mercier *et al.*, 2015). However, the number of DSBs initially formed greatly exceeds the final number of crossovers in *A. thaliana*, *Zea mays*, and possible many other plant species *(Giraut et al, 2011; Ferdous et al, 2012; Lu et al, 2012; Sun et al, 2012; Yang et al, 2012; Drouaud et al, 2013; Wijnker et al, 2013; Qi et al, 2014; Sidhu et al, 2015).*

In plants, more than 80 genes are known to play a role in meiosis (Mercier *et al*, 2015). The analyses of meiotic mutants over the last 30 years have provided detailed insights into the complex mechanisms regulating crossover formation (Figure 2). According to the current model, DSBs are formed by the SPO11 protein (Grelon *et al*, 2001; Stacey *et al*, 2006; Hartung *et al*, 2007) and the broken ends are further resected to generate longer 3'-OH single-stranded DNA. Subsequently, these are bound by the RPA (replication protein A) proteins and loaded by the recombinases RAD51 and DMC1 for homology search and heteroduplex formation (Couteau *et al*, 1999; Osman *et al*, 2009; Da Ines *et al*, 2012). Crossovers may then be formed via two different pathways, namely class I (ZMM pathway) (Börner *et al*, 2004) and class II (MUS81 pathway) (Chelysheva *et al*, 2005; Berchowitz *et al*, 2007). Besides being controlled by different genes, these two pathways result in interference sensitive or insensitive crossovers, which has an effect on the distance between nearby crossovers on the same chromatid (Berchowitz & Copenhaver, 2010).

However, meiotic recombination patterns are not uniform. Genetic divergence in barley, maize, and *Arabidopsis* accessions was shown to account for great variations in recombination frequency (Gale *et al*, 1970; Säll, 1990; Säll *et al*, 1990; Nilsson & Pelger, 1991; Sidhu *et al*, 2015; Ziolkowski *et al*, 2015, 2017). With *A. thaliana* being an ideal model for plant research, (Ziolkowski *et al*, 2015, 2017) were able to show that both *cis* and *trans* effects modify recombination frequency which lead to the discovery of copy number variation in the meiotic E3 ligase gene *HEI10* underlying a recombination QTL (quantitative trait locus).

Extensive work has demonstrated a role for chromatin and epigenetic information in controlling crossover patterns in plants (Yelina *et al*, 2012, 2015; Choi *et al*, 2013; Habu *et al*, 2015; Shilo *et al*, 2015; Mirouze *et al*, 2012; Melamed-Bessudo & Levy, 2012). For example, the histone variant H2A.Z overlaps with crossover hotspots at gene promoters and its deposition is required for wild type levels of recombination (Choi *et al*, 2013; Shilo *et al*, 2015). Furthermore, acquisition of dense DNA methylation and histone H3K9 dimethylation is sufficient to silence crossover hotspot activity (Yelina *et al*, 2015). These associations reflect the classical observation of high meiotic recombination in plant euchromatin and suppression in heterochromatin (Higgins *et al*, 2012; Choi *et al*, 2013;

Phillips *et al*, 2013; Baker *et al*, 2015). Importantly, it has also emerged that epigenetic information, for example DNA and histone methylation, is polymorphic between natural Arabidopsis accessions (Dubin *et al*, 2015; Kawakatsu *et al*, 2016; Moghaddam *et al*, 2011). Therefore, epigenetic variation is also likely to contribute to modification of meiotic frequency when diverged accessions are crossed (Ziolkowski *et al*, 2015, 2017).
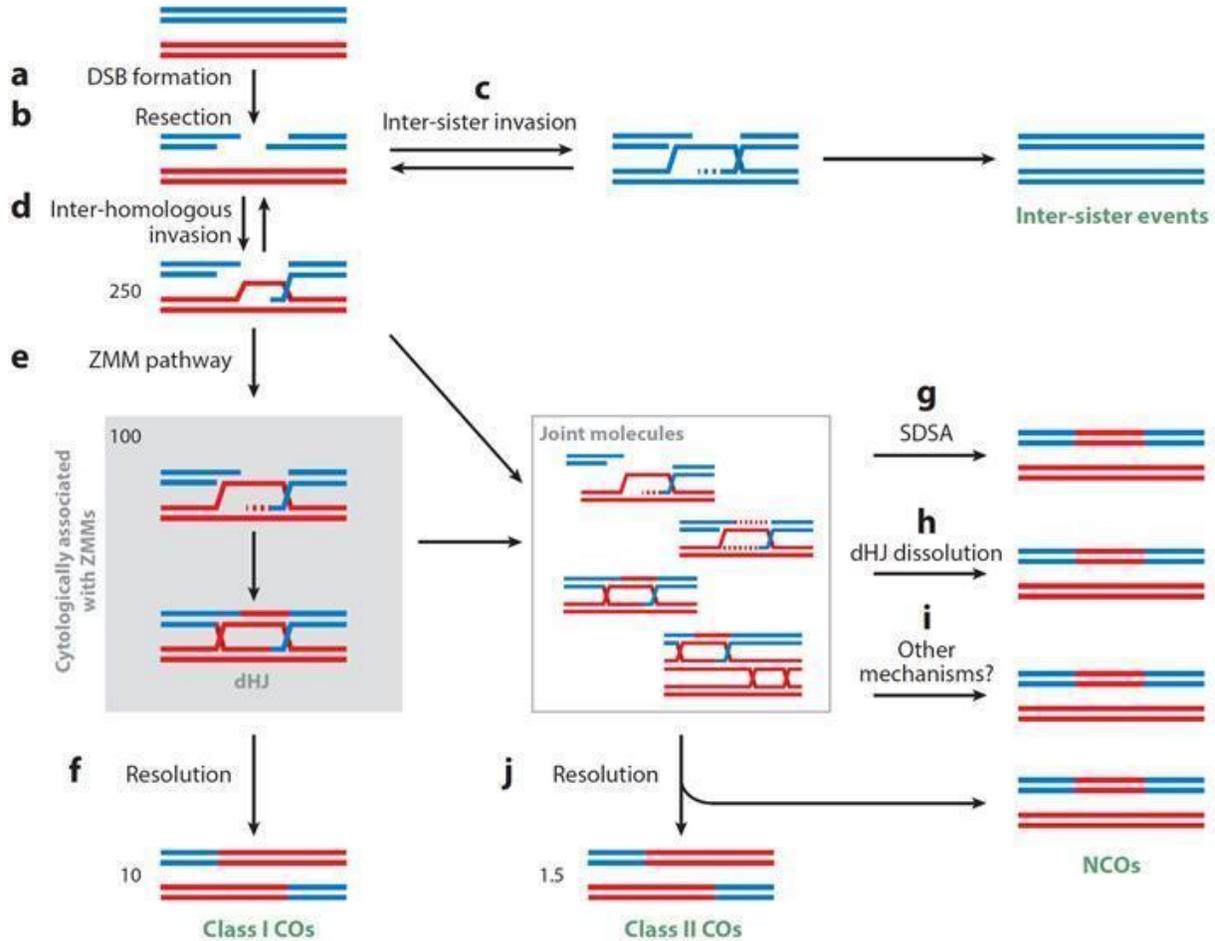
Figure 2 Model of meiotic recombination mechanisms (adopted from Mercier *et al,* 2015). **(a)** Meiotic recombination is initiated by formation of double-strand-breaks (DSBs). **(b)** DSBs are further resected which results in 3' single-stranded DNA overhangs. **(c)** Single-stranded DNA overhangs may invade to sister chromatid for repair. **(d)** Invasion of homologous chromatids by single-stranded DNA. **(e)** Inter-homologous intermediates can be protected by proteins of the ZMM pathway, generating double Holliday junction (dHJ) intermediates. **(f)** Resolution of dHjs as class I crossovers which are sensitive to physical interference. **(g)** Synthesis-dependent strand annealing (SDSA) as an alternative pathway resulting in non-crossovers. **(h)** dHj dissolution as an additional alternative pathway resulting in non-crossovers. **(i)** There is a possibility for other mechanisms to result in non-crossovers. **(j)** A pathway independent of ZMM proteins produces class II crossovers which are insensitive to physical interference.

1.2. Visualization of specific genomic DNA in plants

1.2.1. Staining techniques and development of fluorescent *in situ* hybridization

During the last decades, DNA sequencing technologies have greatly improved our knowledge about the linear sequence order of plant genomes (Arabidopsis Genome Initiative, 2000; International Rice Genome Sequencing Project, 2005; International Brachypodium Initiative, 2010; Dohm *et al*, 2014), and recent chromatin conformation capture approaches are beginning to reveal how plant genomes are spatially organized within the average nucleus (Mascher *et al*, 2017; Feng *et al*, 2014b; Grob *et al*, 2014). However, the spatial organization of plant genomes within the nucleus was also extensively studied by classical microscopy approaches. Visualization of genomic DNA and chromatin in plant nuclei allows us to directly interrogate the spatial organization therein. Classical approaches to visualize chromatin rely on dyes such as Giemsa-stain, aceto-carmin, and aceto-orcein (Lam *et al*, 2004). Such chromatin staining methods are DNA sequence-unspecific but achieve reproducible staining patterns which enable karyotyping and detection of large genome rearrangements (Lam *et al*, 2004). The introduction of 4',6-diamidino-2-phenylindole (DAPI) as a fluorescent dye that specifically associates with the minor groove of double-stranded DNA improved staining intensities by a factor of 20 (Barcellona *et al*, 1990; Kubista *et al*, 1987). Furthermore, DAPI preferentially binds to AT-rich DNA sequences which also happen to be enriched in heterochromatic regions (Lam *et al*, 2004).

Another technique that allows visualization of specific DNA sequences is fluorescent *in situ* hybridization (FISH) which is based on fluorescent dyes conjugated to DNA probes. FISH is a well-established tool to map DNA sequences on chromosomes and visualize spatial DNA arrangements within nuclei (Aliyeva-Schnorr *et al*, 2015; Lysak *et al*, 2001). However, it relies on fixed tissue samples and cannot be used to visualize dynamic processes in living cells. FISH also requires cell fixation and a DNA denaturation step which may result in an altered chromatin structure (Boettiger *et al*, 2016; Kozubek *et al*, 2000).

1.2.2. Live cell imaging of specific DNA sequences in plants

In contrast to classical and fluorescence-based cytological methods, live cell imaging of specific DNA sequences allows us to investigate spatio-temporal dynamics of chromosomes. As already mentioned above (see 1.1.2), chromatin is not rigidly fixed within the nucleus but is dynamic, and chromosomal regions may change their location as a result of or prerequisite for transcriptional activation and silencing, as well as response to environmental stimuli.

Live cell imaging of specific genomic loci in plants has been achieved first by the application of a directly repeated *lac* operator sequence and its detection with a GFP-tagged lac repressor protein (Kato & Lam, 2001). However, this system is based on the random insertion of an exogenous sequence into the genome. Live imaging of endogenous genomic regions became possible with the discovery of Cys2-His2 zinc-finger proteins from *Xenopus (Miller et al, 1985)*. Individual zinc-fingers are composed of 30 amino acids (Beerli & Barbas, 2002) where each unit can be engineered to recognize a nucleotide triplet (Pavletich & Pabo, 1991; Klug, 1993). Multiple zinc-fingers can be arranged in tandem as a zinc-finger protein (ZFP) to recognize specific DNA sequences (Choo *et al*, 1994; Pabo *et al*, 2001). Fluorescent protein-tagged ZFPs were first used in *A. thaliana* and *Mus musculus* for live cell imaging of repetitive DNA (Lindhout *et al*, 2007). Despite numerous studies where ZFPs were engineered for genome editing (Beumer *et al*, 2006; Bibikova *et al*, 2003; Zhang *et al*, 2010), the potential of this technology has not been fully exploited yet for imaging, probably due to its difficult design.

Transcription activator-like effectors (TALEs) were first discovered in the plant pathogenic bacteria *Xanthomonas* and can be engineered to bind specific DNA sequences through tandemly arranged 33 - 35 amino acid repeats where each repeat binds to one nucleotide (Boch *et al*, 2009). While TALEs fused with fluorescent proteins were successfully used for live cell imaging in cultured cells and live organisms (Ma *et al*, 2013; Yuan *et al*, 2014), their application in plants was only recently shown (Fujimoto *et al*, 2016).

The discovery of the type II clustered regularly interspaced short palindromic repeats (CRISPR)-associated protein 9 (Cas9) system derived from *Streptococcus pyogenes* revolutionized the field of targeted genome editing in eukaryotes (Jinek *et al*, 2012).

CRISPR-Cas9-based genome engineering has become a routine technology in many plant species (Pacher & Puchta, 2017). Recently, nuclease-deficient derivatives of Cas9 (dCas9) in combination with transcription activators (EDLL and TAL effectors) and repressors (SRDX repression domain, CRISPR interference) were used to modify gene expression in many model organisms including plants (Piatek *et al*, 2015; Qi *et al*, 2013). Furthermore, by fusing dCas9 with GFP, CRISPR-dCas9 was used to label genomic loci in live mammalian cells (Chen *et al*, 2013; Anton *et al*, 2014). Multicolour CRISPR-dCas9 imaging was enabled through utilization of dCas9 orthologues from different bacteria species, like *Neisseria meningitidis* (NmCas9), *Streptococcus thermophilus* (St1Cas9), and *Staphylococcus aureus* (Ma *et al*, 2015; Chen *et al*, 2016). The discovery of the Cas9-like activities of the Cpf1 protein derived from *Acidaminococcus* and *Lachnospiraceae* (Zetsche *et al*, 2015) may further expand the palette of engineered proteins for live cell imaging of DNA. In addition to visualization of DNA, CRISPR-dCas9 was also used to visualize RNA in living cells through DNA oligonucleotides that would hybridize with the target RNA and supply the protospacer adjacent motif (PAM) required for binding (Nelles *et al*, 2016; O'Connell *et al*, 2014).
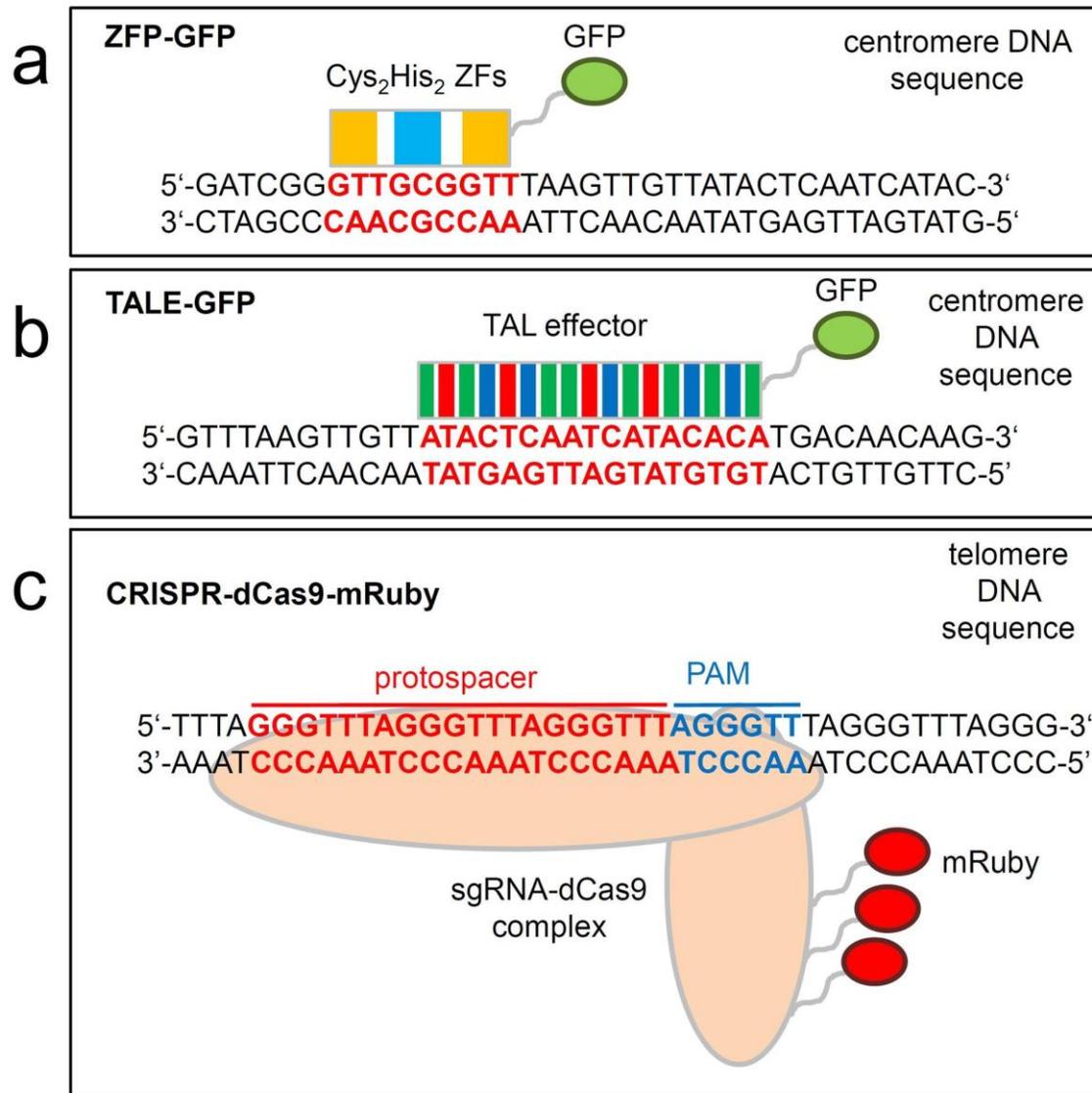
Figure 3 Overview of engineered proteins used for live cell imaging in plants.
(a) A zinc-finger protein (ZFP) fused with GFP was engineered to recognize a 9-bp sequence within the *A. thaliana* centromeric DNA sequence pAL1 (Lindhout *et al*, 2007). Individual Cys_2-His_2 zinc-fingers (ZFs, shown as orange or blue rectangles) recognize 3-bp triplets of the target sequence (highlighted in red). **(b)** Transcription-activator-like (TAL) effector fused with GFP was engineered to recognize a 17-bp sequence within the *A. thaliana* centromere DNA sequence pAL1 (Fujimoto *et al*, 2016). Individual units that specifically recognize certain nucleotides (highlighted in red) are shown as orange, blue, and magenta rectangles. **(c)** CRISPR-dCas9 derived from *S. aureus* fused to three copies of mRuby forms a complex with a single guide (sg) RNA containing a sequence complementary to telomeric DNA. Protospacer motif is highlighted in red. Protospacer adjacent motif (PAM) which is required by *S. aureus* dCas9 (NNGRRT) is highlighted in blue.

1.3. Single-cell analysis: From rare microbes and human cancer cells to meiosis research in plants

1.3.1. Applications of single-cell genome sequencing

Single-cell analysis differs from bulk-cell or cell-population analysis in a way that it reveals heterogeneity between individual cells in a complex biological system such as a developing embryo, a human tumor, a microbial ecosystems, and different plant tissues. With the advance of single-cell analysis approaches during the last decades, it became possible to address new questions in microbial research, human genetics, and plant research. For example, it was estimated that approximately 99% of all microbial species cannot be grown in pure culture which used to leave their genomes inaccessible (Kalisky *et al*, 2011; Blainey & Quake, 2014). Through the advance of single-cell genome sequencing, several uncultivated microbes were successfully sequenced. These include TM7, a candidate phylum with environmental and clinical relevance (Marcy *et al*, 2007), marine microbes (Rodrigue *et al*, 2009; Woyke *et al*, 2009), an insect symbiont (Woyke *et al*, 2010), organisms from a complex microbial community such as cattle rumen (Hess *et al*, 2011), and the first single-cell archaeal genome of *Nitrosoarchaeum limnia (Blainey et al, 2011)*. In addition, single-cell sequencing was used to unravel the sexual origin of heterokaryosis in arbuscular mycorrhizal fungi (Ropars *et al*, 2016).

In human genetics, single-cell analysis holds promise to reveal the complexity of tumor heterogeneity, where it is unknown which tumor cells are responsible for the respective phenotype, and which could improve tumor treatment (Wang *et al*, 2014). Furthermore, it also enables to understand important somatic variation caused by gradual accumulation of mutations through errors in cell division and mobile genetic elements which is associated with aging and cancer (Lynch, 2010).

In plants research, single-cell approaches offer exciting new possibilities. For example, single-chromosome sequencing, which is a related approach, was used to improve the reference sequence of wheat chromosome 3B (Cápal *et al*, 2015) and to identify the chromosomal location of a transgene in wheat (Cápal *et al*, 2016). In addition, single-cell sequencing can be used in meiosis research to measure meiotic recombination, investigate gene conversion, determine segregation ratios, and provide a reference

genetic map (Li *et al*, 2015; Dreissig *et al*, 2015). Its greatest advantage in plant meiosis research might be to overcome the sampling issue related to other methodologies such as microscopy and segregating population analysis. Meiotic recombination events are quite rare with typically only 1 - 2 events per chromosome pair per meiosis.

There are numerous methods to measure meiotic recombination in plants, including molecular markers (Salomé *et al*, 2012), cytological visualization of crossovers (Sybenga, 1966; Anderson *et al*, 2003; Phillips *et al*, 2013), tetrad analysis (Copenhaver *et al*, 2000), fluorescent protein-tagged loci expressed in pollen (Berchowitz & Copenhaver, 2008; Francis *et al*, 2007; Yelina *et al*, 2013), and several pollen genotyping approaches (Drouaud & Mézard, 2011; Khademian *et al*, 2013; Li *et al*, 2015). Although these methods have been successfully used to characterize recombination patterns and improve our understanding of meiosis, each of them has its specific advantages and disadvantages. The analysis of recombination through genotyping of a segregating population is laborious and very challenging for some plant species. Cytological analysis of recombination is more widespread and applicable to many plant species, yet its resolution is lower compared to sequence-based approaches and the analysis is demanding in terms of time and experience. Tetrad analysis combined with fluorescence markers is a very powerful high-throughput approach but requires the integration of reporter transgenes and is so far limited to the model species *Arabidopsis* (Berchowitz & Copenhaver, 2008; Francis *et al*, 2007; Yelina *et al*, 2013),

As it was previously shown with maize microspores, single-cell sequencing is capable of reliably measuring meiotic recombination (Li *et al*, 2015). With further advances in the field of genomics and decreasing costs of sequencing, these single-cell approaches might enable us to analyse large numbers of meiotic cells at unprecedented resolution. It is therefore reasonable to think that single-cell genome sequencing will complement and even outperform microscopy- or population-based approaches in terms of analysing meiotic recombination.

1.3.2. Technical challenges of single-cell sequencing

There are four major technical challenges associated with single-cell genome sequencing, (i) physical isolation of individual cells from a tissue or population, (ii) amplification of the cellular DNA to acquire sufficient amounts for downstream analyses, (iii) analysis of the amplified DNA in a cost-efficient way to address the hypothesis of the study, and (iv) interpretation of the data in the context of biases and errors which might arise during the previous steps (Gawad *et al*, 2016).

Several approaches for single-cell isolation were successfully used, including microfluidic flow-sorting (Marcy *et al*, 2007; Wang *et al*, 2012), fluorescence activated cell sorting (FACS, (Cápal *et al*, 2015; Ropars *et al*, 2016; Rodrigue *et al*, 2009; Woyke *et al*, 2009)), micromanipulation (Woyke *et al*, 2010; Houben *et al*, 1996; Matsunaga *et al*, 1999), and optical tweezing (Blainey *et al*, 2011). Different tissues or different organisms can pose their own unique challenges in terms of cell preparation. For example, plant pollen are composed of rigid cell walls (Shi *et al*, 2015; Goss, 1968) that may require enzymatic or mechanical disruption of the cell wall (Chen *et al*, 2008; De Storme & Geelen, 2011). At the same time, their nuclei need to remain intact for isolation through fluorescence activated cell sorting (FACS) and subsequent downstream analyses.

The second step is amplification of the DNA of a single cell for downstream analyses. Most cells do only contain picogram amounts of DNA, depending on genome size, which is insufficient for direct downstream analyses such as whole-genome sequencing. In general, the aim is to amplify cellular DNA up to microgram amounts by simultaneously minimizing the introduction of artefacts, such as amplification bias, locus loss, mutations, and sequence chimaeras (Gawad *et al*, 2016). Therefore, several whole-genome amplification (WGA) methods were developed that each have their advantages and disadvantages which need to be considered carefully for different experiments (Figure 4). At first, pure PCR-based methods were developed which utilized degenerate or random oligonucleotide priming (Telenius *et al*, 1992; Zhang *et al*, 1992). These methods are associated with high locus loss rates due to differences in density of common sequences and variability in PCR efficiency between loci (Gawad *et al*, 2016). Secondly, WGA methods based on multiple-displacement amplification (MDA) were developed that rely on quasi-random priming and isothermal amplification using the Φ29 DNA polymerase

(Dean *et al*, 2001; Zhang *et al*, 2001). These methods result in higher genome coverage and fewer mutations due to increased fidelity of the Φ29 DNA polymerase (de Bourcy *et al*, 2014). However, it also introduces amplification bias in terms of overrepresentation of those loci which get amplified first (de Bourcy *et al*, 2014). Finally, hybrid methods such as multiple annealing and loop-based amplification cycles (MALBAC) and displacement-DOP-PCR (also known as PicoPLEX) were developed that both use limited isothermal amplification followed by PCR amplification of the products of the first step (Zong *et al*, 2012; Langmore, 2002). These methods were conceptualized in order to overcome the low coverage of PCR-based methods and the low uniformity of MDA-based methods. In a representative comparison using serial dilutions of *E. coli* DNA as well as single bacterial cells, both MDA and hybrid methods were successfully used to amplify DNA from single cells (de Bourcy *et al*, 2014). However, both methods amplified a significant amount of contaminant DNA when carried out in microlitre volumes in standard reaction tubes. This contamination was largely avoided by carrying out the same reactions in nanolitre volumes in a microfluidic system. This illustrates another principle of single-cell WGA approaches, which is that smaller reaction volumes result in higher amplification efficiencies due to higher target DNA concentration and lower amounts of contamination (Marcy *et al*, 2007; Wang *et al*, 2012; Gole *et al*, 2013).

After successful DNA amplification, there are several options available how to analyse the amplified DNA to address the hypothesis of the study. For example, one may consider to carry out multiple PCR reactions to analyse loci of interest (as it was done in the current study), perform exome-capture to focus on protein-coding regions (Leung *et al*, 2015), or to sequence the entire genome (Wang *et al*, 2012; Li *et al*, 2015). However, it is important to consider the costs of each method in relation to the question that needs to be addressed (Figure 5). For example, it might be interesting to analyse meiotic recombination between two loci in thousands of cells, where it would be of advantage to use allele specific PCR. On the other hand, one would prefer whole-genome sequencing if interested in fine scale analysis of the meiotic recombination landscape.

Finally, it is of great importance to interpret the data in the context of biases and errors which might arise during the previous steps. It is therefore important to consider including controls in single-cell experiments, such as purely homozygous parental cells when

analysing meiotic recombination. Homozygous cells should not show different alleles and can therefore be used to interrogate errors that could arise during DNA amplification. In addition, cell isolation through microfluidics or flow cytometry can be biased by selection of cells based on size or viability. It is therefore important to compare results obtained from single-cell experiments with those obtained through different methods, such as the analysis of meiotic recombination in single F_1 pollen nuclei versus a double haploid (DH) population of the same genotype.
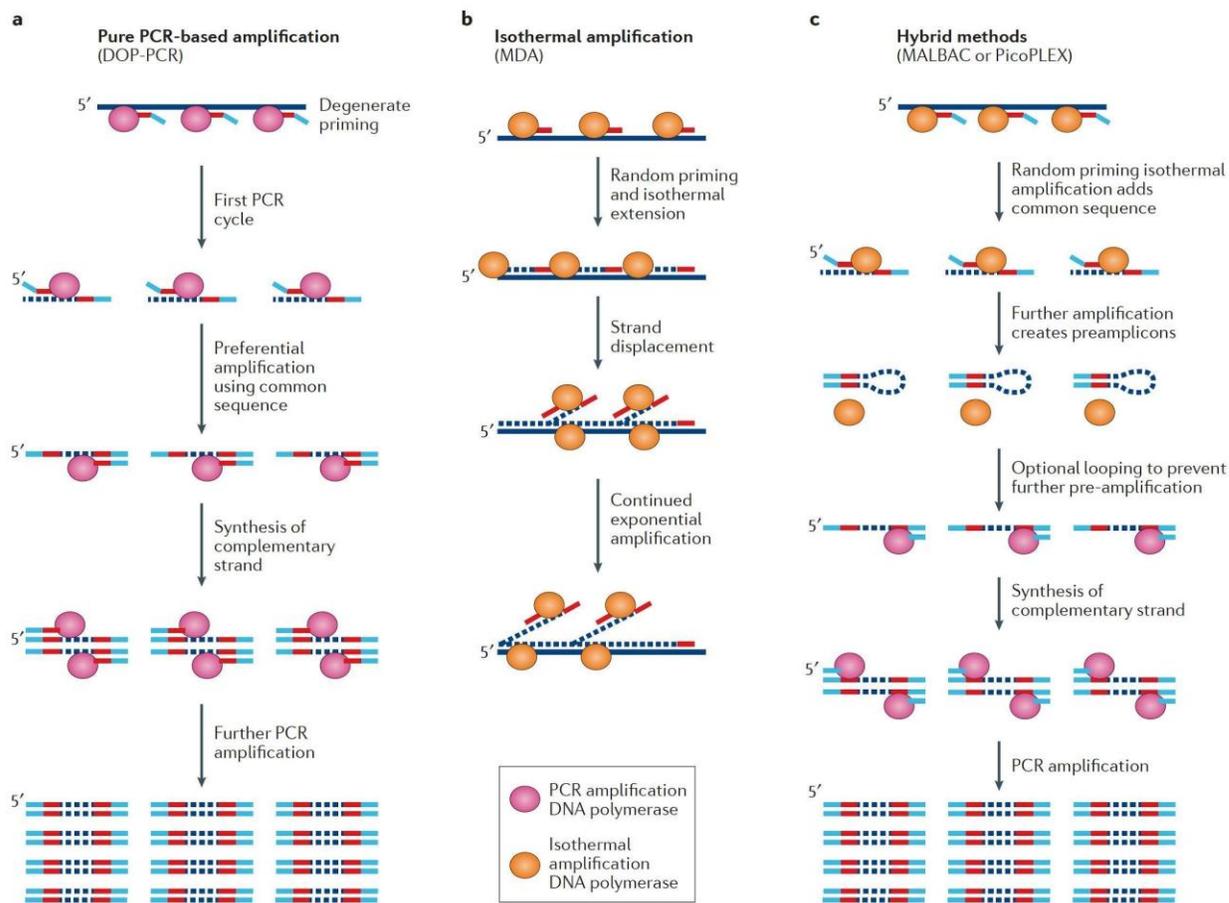
Figure 4 Overview of the three main whole-genome amplification methods (adopted from Gawad *et al*, 2016). **(a)** PCR-based method using degenerate oligonucleotide primed PCR (DOP-PCR) using random priming followed by PCR amplification. Specific sites in the genome are preferentially amplified which results in low genome coverage but better uniformity of amplification. **(b)** Isothermal multiple displacement amplification (MDA) using random priming combined with a Φ29 DNA polymerase with high processivity and strand displacement activity. This methods has high genome coverage but is less uniform across the genome. **(c)** Hybrid methods such as multiple annealing and looping based amplification cycles (MALBAC) and PicoPLEX using an initial isothermal amplification, during which common sequences are added to the amplicons, followed by PCR amplification using those sequences. These methods have intermediate genome coverage and uniformity compared to pure DOP-PCR and MDA.
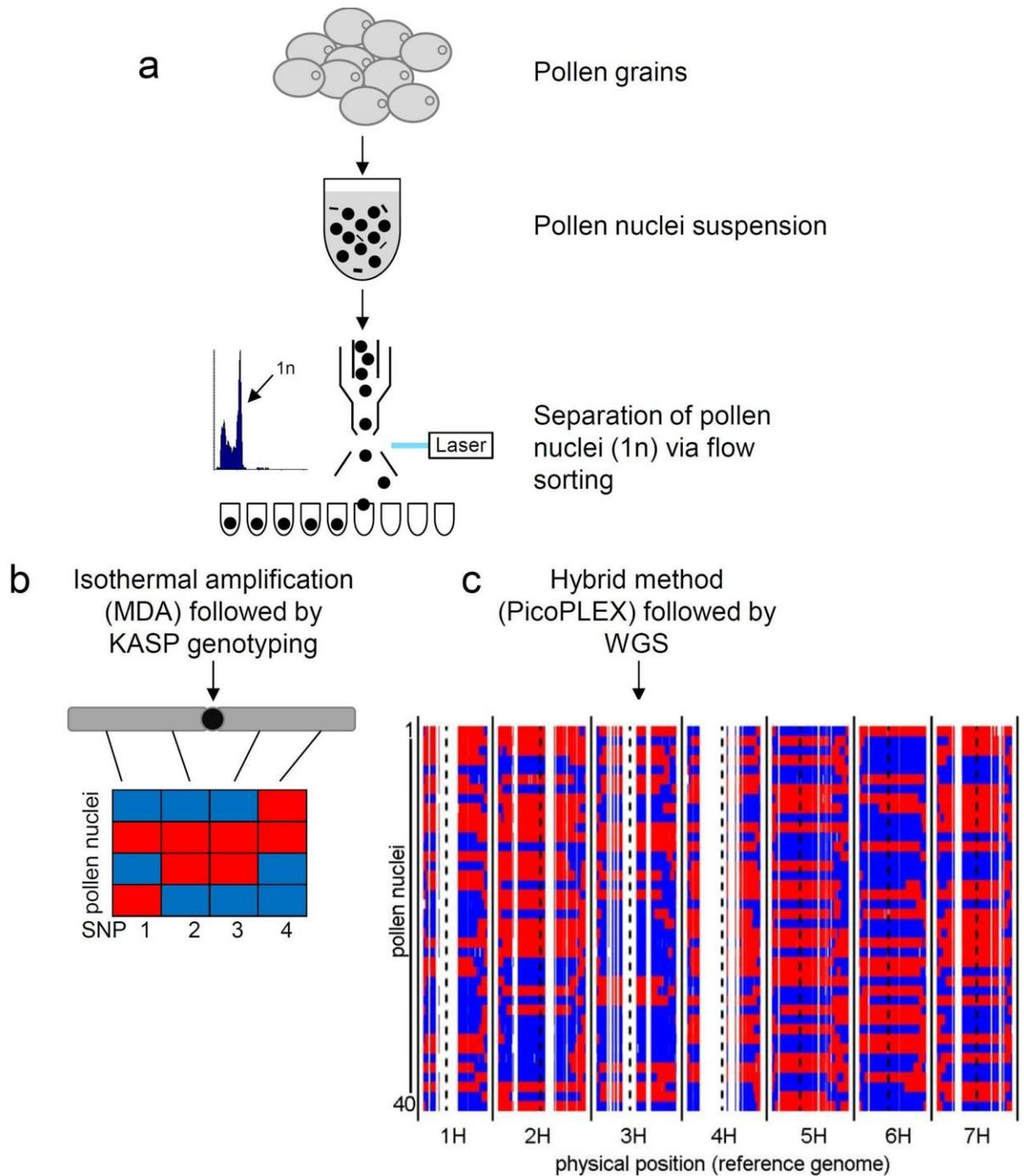
Figure 5 Overview of two different single-cell analysis approaches.
(a) Pollen grains are disrupted by metallic beads which releases nuclei. Pollen nuclei are stained by DAPI and flow cytometry is applied to sort individual haploid pollen nuclei into separate reaction tubes of a microwell plate for whole-genome-amplification. **(b)** Isothermal multiple displacement amplification (MDA) is applied on individual pollen nuclei followed by KASP genotyping to score recombination frequency in one or more defined chromosomal intervals. This model shows a chromosome (grey) and its centromere (black) with 4 SNPs with known physical position (thin black lines). Recombination events

are then measured by allele-specific KASP markers (blue vs red) in individual pollen nuclei. This approach offers relative ease of data analysis but focuses on specific chromosomal intervals of interest. **(c)** PicoPLEX amplification, which is a combination of an initial isothermal amplification followed by PCR re-amplification and library preparation for Illumina sequencing, is followed by Illumina paired-end sequencing to generate a genome-wide recombination map of individual pollen nuclei. Recombination events are visualized as allelic changes (blue vs red) along all seven barley chromosomes. This method offers high reproducibility and resolution of meiotic recombination at megabase scale across the whole genome.

2. Aims of the study

In the current study, my aim was to advance single-cell analysis in plants by developing novel methodologies that would pave the way for future work in the field of basic and applied plant research.

The first part of the study is focused on meiotic recombination and its implications for plant breeding. A major focus of plant meiosis research is on modifying recombination patterns in crops to unlock genetic diversity and aid crop improvement through breeding. Many of these attempts require efficient methods to measure meiotic recombination in mutant backgrounds, in response to environmental stimuli, in chemically altered epigenetic contexts, and divergent accessions. Therefore, my aim was to develop a new methodology to measure meiotic recombination based on genotyping or sequencing of single pollen nuclei and provide evidence for its performance compared to standard methods.

The second part of this study is focused on the visualization of defined genomic regions in live cells for basic plant research, but presumably also future applied research in plants. Interphase chromatin dynamics might be a result of or prerequisite for transcriptional activation of genes in response to environmental stimuli or developmental processes. Our means to visualize specific genomic regions in live plant cells in order to observe these dynamics are still limited. I was therefore aiming to take advantage of the recently discovered CRISPR-Cas9 system and establish a CRISPR-Cas9-based live cell imaging method in plants.

Together these new methodologies will enable to better understand meiotic recombination and interphase chromatin dynamics in basic and applied research.

3. Publications

3.1. Measuring meiotic crossovers via multi-locus genotyping of single pollen grains in barley

Dreissig, S. et al., 2015. Measuring meiotic crossovers via multi-locus genotyping of single pollen grains in barley. PloS one, 10(9), p.e0137677.

Measuring Meiotic Crossovers via Multi-Locus Genotyping of Single Pollen Grains in Barley

Steven Dreissig[1], Jörg Fuchs[1], Petr Cápal[2], Nicola Kettles[3], Ed Byrne[3], Andreas Houben[1]*

1 Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstrasse 3, 06466 Stadt Seeland, Germany, 2 Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, Olomouc CZ-78371, Czech Republic, 3 KWS-UK Ltd, 56 Church Street, Thriplow, Hertfordshire, SG8 7RE, United Kingdom

* houben@ipk-gatersleben.de

Abstract

The detection of meiotic crossovers in crop plants currently relies on scoring DNA markers in a segregating population or cytological visualization. We investigated the feasibility of using flow-sorted haploid nuclei, Phi29 DNA polymerase-based whole-genome-amplification (WGA) and multi-locus KASP-genotyping to measure meiotic crossovers in individual barley pollen grains. To demonstrate the proof of concept, we used 24 gene-based physically mapped single nucleotide polymorphisms to genotype the WGA products of 50 single pollen nuclei. The number of crossovers per chromosome, recombination frequencies along chromosome 3H and segregation distortion were analysed and compared to a doubled haploid (DH) population of the same genotype. The number of crossovers and chromosome wide recombination frequencies show that this approach is able to produce results that resemble those obtained from other methods in a biologically meaningful way. Only the segregation distortion was found to be lower in the pollen population than in DH plants.

Introduction

Meiotic recombination is the primary mechanism of generating novel allelic combinations and introducing genetic diversity. In barley (*Hordeum vulgare* L.), as well as in many other crops, recombination frequencies are elevated in distal gene-rich chromosomal regions. Nevertheless, 24.7% of the total barley gene content is located in low recombining regions [1] representing an untapped resource which is unavailable for plant breeding [2]. Hence, strategies to modulate the recombination frequency along chromosomes are needed. The ability to induce an increase in meiotic recombination is so far limited to the model species *Arabidopsis thaliana* via a mutation of the FANCM helicase [3]. In barley, Higgins et al. [4] demonstrated a shift of meiotic crossovers towards interstitial and proximal regions at higher temperatures during meiosis.

There are different ways of monitoring meiotic crossovers in plants. They can be identified using molecular markers in a segregating population [5], or alternatively the frequency and distribution of crossovers can be visualized by cytological means like analysis of pairing configurations [6] or immunolabeling of proteins involved in meiotic recombination such as the barley

MutL homologue (HvMLH3) [7]. One limitation of using microscopy-based methods is that the sites of recombination events can only be resolved at the chromosomal level. Another limitation is an uncertainty about perfect agreement between protein localization and crossover. Other tools for efficient determination of recombination events such as the tetrad analysis based on the quartet (qrt) mutation are currently only available for *Arabidopsis thaliana* [8].

In human and livestock genetics, recombination analysis using meiotic gametophytes was developed more than 20 years ago for the high-resolution mapping of recombination sites. In plants, the idea of analysing pollen grains has already brought forward a number of studies. Petersen et al. [9] extracted DNA from single barley and rye pollen grains for PCR amplification and subsequent sequencing of high and single copy genes. Chen et al. [10] developed a method using pollen grains of several plant species for molecular analysis utilizing randomly amplified polymorphic DNA and simple sequence repeat markers. The introduction of whole-genome-amplification (WGA) methods, such as primer extension pre-amplification, enabled Aziz and Sauve [11] to further increase the amount of information gained from single pollen grains. However, other WGA methods based on isothermal amplification via the Phi29 DNA polymerase hold the potential to enable the analysis of hundreds to thousands of markers in a single cell [12].

In the current study, we describe a strategy to perform a parallel analysis of individual haploid nuclei derived from pollen grains by utilizing fluorescence activated cell sorting (FACS) coupled with Phi29 DNA polymerase-based whole-genome-amplification (WGA) and multilocus KASP genotyping. The meiotic crossover measurements were compared to data obtained by different methods in comparable genetic environments.

Materials and Methods

Plant material and isolation of pollen nuclei and genomic DNA

Pollen grains were collected from an F_1 plant of the barley cultivars Morex x Barke (*Hordeum vulgare* L.). Mature anthers of 20 flowers were collected in a 1.5 ml Eppendorf tube using forceps. Afterwards, 300 µl of ddH$_2$O were added and the suspension was vortexed for approximately 30 sec. The suspension was shaken at 1500 rpm for 10 min at room temperature to release all pollen grains. Afterwards, the pollen suspension was centrifuged for 5 min at 13,000 rpm and all empty anthers were manually removed using forceps. After centrifugation for 5 min at 13,000 rpm the supernatant was removed and the pollen pellet was resuspended in 100 µl Galbraith buffer (45 mM MgCl$_2$, 30 mM sodium citrate, 20 mM MOPS, 0.1% Triton-X100, pH to 7.0; [13]) and transferred into a 2.0 ml Eppendorf tube containing two metallic beads of 6 mm in diameter (Intec GmbH) as described in [14]. The pollen-bead mixture was centrifuged for 5 min at 13,000 rpm prior to homogenization at 30 Hz for 40 seconds using a MM 400 ball mill (Retsch). After homogenization, another 500 µl Galbraith buffer were added and the suspension was filtered through a 30 µm filter (Sysmex-Partec). For the purpose of providing genomic DNA for marker testing, genomic DNA was extracted from leaf tissue using the DNeasy Plant Mini kit (Qiagen) and measured using nanodrop (Peqlab).

FACS-based purification of single haploid nuclei and whole-genome-amplification

The nuclei suspension was stained with 4',6-diamidino-2-phenylindole (DAPI; 1.5 µg/ml) and single 1C nuclei were sorted using a BD FACSAria IIu (BD Biosciences) flow-sorter into individual wells of a 384-microwell plate containing 2 µl lysis solution (0.5 µl lysis buffer composed of 400 mM KOH, 100 mM DTT, 10 mM EDTA; [15], 0.5 µl ddH$_2$O, 1 µl sample buffer

(Genomiphi V2, GE Healthcare)) for whole-genome-amplification. Note, in contrast to the manufacturer's protocol the sample buffer containing random primers for whole-genome-amplification was added to lysis solution. Whole-genome-amplification was carried out using the Genomiphi V2 kit (GE Healthcare) according to the manufacturer's protocol with the following modifications: Nuclei lysis and DNA denaturation was conducted by incubation at 65°C for 3 min in 2 µl lysis solution. The lysis solution was neutralized by adding 0.5 µl neutralisation buffer (666 mM Tris-HCl, 250 mM HCl; [15]). Afterwards, a master mix composed of 3.5 µl sample buffer, 4.5 µl reaction buffer and 0.5 µl enzyme mix (all Genomiphi V2, GE Healthcare) per reaction was added and samples were incubated at 30°C for 8 hours followed by inactivation of the enzyme at 65°C for 10 minutes. Subsequently, each sample was diluted with 500 µl ddH$_2$O. The DNA concentration of the WGA products of single pollen nuclei was measured by fluorometric quantitation (Qubit, Life Technologies).

Each sample was subjected to a PCR using primers for the *Ty3/gypsy*-like retroelement *cereba* in order to validate the successful sorting of pollen nuclei into the microwells. The reaction volume of the *cereba* amplification was 10 µl containing 5 µl WGA product, 1x PCR buffer (Qiagen), 0.2 mM dNTPs (Bioline), 1x Q-solution (Qiagen), 0.6 µM of each primer and 0.02 units Taq DNA polymerase (Qiagen). The following thermal cycling conditions were used: DNA polymerase activation: 3 min at 95°C; denaturation: 30 sec at 95°C; annealing: 30 sec at 60°C; extension: 30 sec at 72°C; final extension: 10 min at 72°C; 30 cycles in total. The *cereba*-positive samples were further analysed with 8 chromosome 3H-specific primers to quantify the efficiency of the whole-genome-amplification (S1 Table). These primer pairs were targeting single copy sequences to test if the WGA was able to amplify unique sequences. The reaction volume of the 3H-specific amplification was 10 µl containing 5 µl WGA product, 1x PCR buffer (Qiagen), 0.2 mM dNTPs (Bioline), 1x Q-solution, 0.3 µM of each primer and 0.02 units Taq DNA polymerase (Qiagen). The following thermal cycling conditions were used: DNA polymerase activation: 3 min at 95°C; denaturation: 30 sec at 95°C; annealing: 30 sec at 65°C, reduced by 1°C for 9 cycles; extension: 30 sec at 72°C; 25 cycles at final annealing temperature.

KASP-genotyping

A set of 24 chromosome 3H-specific single nucleotide polymorphisms (SNPs) [16] based on the current barley genome sequence assembly [17] was chosen and converted into KASP markers (LGC Genomics, S2 Table). Thermal cycling conditions were adopted from Mirouze et al. [18] and end-point signals were read out on a BioRad iQ5 cycler at 30°C. Genomic DNA from cv. Morex, cv. Barke and Morex x Barke F$_1$ plants were genotyped in parallel. Additionally, replicate single pollen nuclei of the cultivar Morex were subjected to whole-genome-amplification and subsequent genotyping to act as a positive control against amplification errors. This was done in order to test the possibility of false allele calling due to inaccurate WGA or contamination. Allele calling was done manually by plotting relative fluorescence values of FAM and HEX against each other. Heterozygous signals were discarded as genotyping errors since we expect haploid nuclei to only provide homozygous signals.

Analysis of segregation distortion loci and crossovers

To test for segregation distortion we conducted a χ^2-test assuming an expected segregation ratio of 1:1 for each marker. Segregation distortion loci (SDL) were identified by significant deviation from the expected ratio of 1:1 ($P < 0.05$). Crossovers were detected by visualizing our marker data using flapjack [19] by identifying allele calls for which there was a switch from allele A (Morex) to allele B (Barke) and vice versa. The physical position of each marker on barley chromosome 3H was derived from the barley genome sequence assembly [17] thus enabling

us to count crossovers without constructing a linkage map. The recombination frequency between two adjacent marker pairs was measured as the proportion of crossovers to no-crossovers. Marker pairs within a sample involving missing data points were omitted from the analysis. The number of crossovers was normalized according to the following calculation:

$$\text{CO ratio} = \frac{n(CO)}{n(no-CO) + n(CO)} \tag{1}$$

where n(CO) is the number of crossovers and n(no-CO) is the number of no-crossovers.

Similarly, the number of no-crossovers was normalized according to the following calculation:

$$\text{no-CO ratio} = \frac{n(no-CO)}{n(no-CO) + n(CO)} \tag{2}$$

The ratio of (1) to (2) enabled us to calculate recombination frequencies for each marker pair in a comparable manner. To compare our data to the Morex x Barke DH population data, we used raw genotyping-by-sequencing data (wheat.pw.usda.gov, S3 Table) and extracted the physical map position for each marker [17]. Segregation distortion and crossover analysis of the Morex x Barke DH population was performed as described above. A two-tailed unpaired Student's t-test was performed to compare the average number of crossovers in both populations and the distribution of the number of crossovers was compared using a χ^2-test assuming the Morex x Barke DH crossover data as expected values.

Results and Discussion

Whole-genome-amplification of single haploid nuclei

In order to develop a strategy for the high-throughput analysis of meiotic recombination events in barley pollen, we first needed to prove the successful extraction, amplification and genotyping of single pollen DNA. To overcome problems associated with the rigid cell wall of pollen grains described by Chen et al. [10], we selected a novel approach to isolate haploid nuclei suitable for flow-sorting.

Isolated haploid nuclei from pollen grains were individually sorted via a FACS-based approach into individual wells of a 384-microwell plate (Fig 1). After nuclei lysis and DNA denaturation, whole-genome-amplification was performed using Phi29 DNA polymerase. Each reaction yielded 1 to 3 µg of DNA consisting of barley-specific products and likely also unspecific products as expected from whole-genome-amplification via Phi29 DNA polymerase [20]. The products of 192 single-nuclei whole genome amplification (WGA) reactions were analysed by PCR for the presence of the barley high copy *Ty3/gypsy*-like retroelement *cereba* to confirm the successful sorting of nuclei into the individual microwells. From a total of 192 samples, 168 contained PCR amplified barley DNA giving an accuracy of our FACS approach of 87.5%. To preselect samples for further genotyping, we used PCR to amplify eight 3H-specific single copy sequences located across both arms of chromosome 3H. Out of the 168 single-nuclei amplifications positively tested for *cereba*, we selected 50 samples which showed successful amplifications of at least 3 single copy sequences from both arms for further genotyping (S1 Fig). To measure meiotic crossovers on chromosome 3H, we selected 24 KASP markers. The suitability of the 24 selected chromosome 3H-specific KASP markers was confirmed using genomic DNA isolated from leaves of both genotypes. Next, the same set of markers was used to genotype WGA-amplified DNA derived from individual haploid nuclei. The selected 50 nuclei samples revealed an average marker call rate of 70%, indicating that the majority of samples were effectively amplified (Fig 2A). Considering preselection via PCR, we found a weak

Pollen grains

Pollen nuclei suspension

Separation of pollen nuclei
(1n) via flow sorting

Whole-genome-amplification
of single haploid nuclei (1n)

Evalutation of whole-
genome-amplification
efficiency by PCR

Multi-locus genotyping

3H PCR marker

■ positive reaction
■ negative reaction

KASP marker

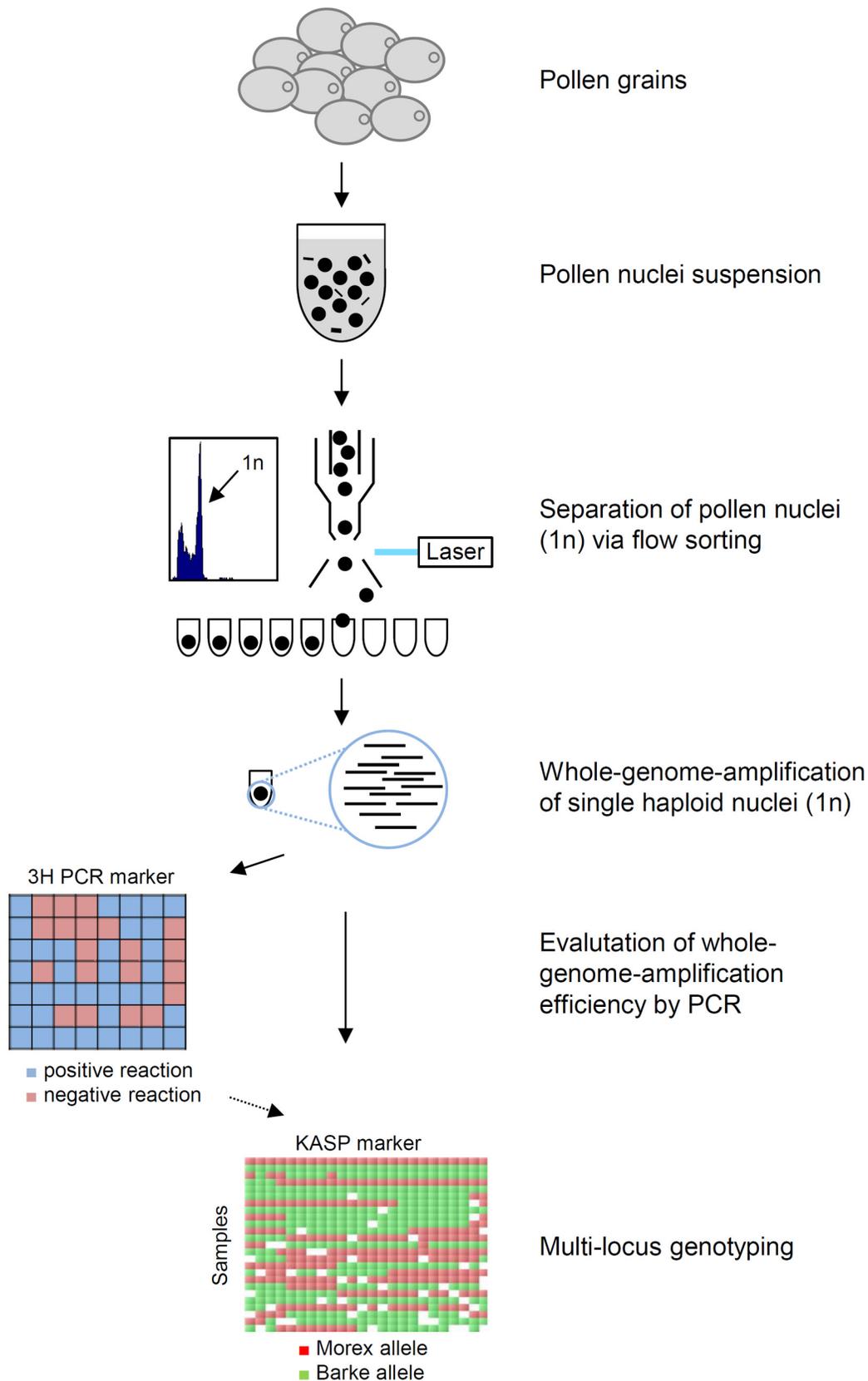Samples

■ Morex allele
■ Barke allele

Fig 1. Scheme of experimental workflow developed in the current study. Haploid nuclei were extracted from pollen grains, separated via flow-sorting and individually subjected to whole-genome-amplification (WGA). High quality samples, evaluated by PCR with chromosome 3H-specific primers, were genotyped using 25 KASP markers to measure crossover frequency and distribution.

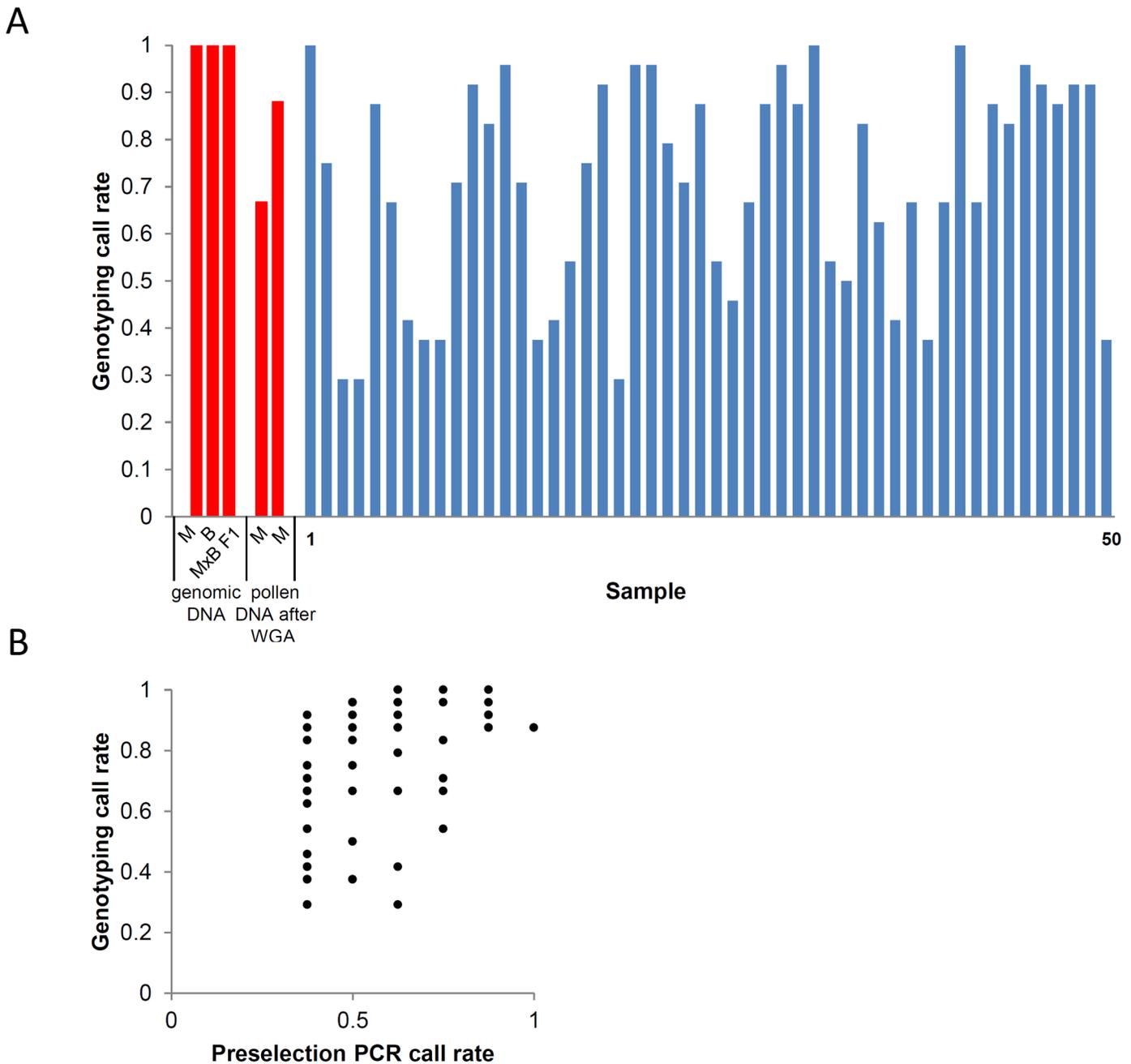doi:10.1371/journal.pone.0137677.g001

Fig 2. KASP genotyping performance. (A) Genotyping call rate of the positive controls across 24 KASP markers is indicated in red. Genomic DNA from Morex, Barke and Morex x Barke F1 plants was used. Two individual nuclei derived from Morex pollen grains were used to test for false allele calling due to whole-genome-amplification. The selected 50 Morex x Barke pollen nuclei samples are indicated in blue showing an average genotyping call rate of 71%. **(B)** Correlation between KASP genotyping call rate and preselection PCR call rate. r = 0.51, r² = 0.26.

doi:10.1371/journal.pone.0137677.g002

but yet positive correlation between final genotyping call rate and preselection PCR call rate (r = 0.51, r² = 0.26, Fig 2B) which indicates the advantage of preselecting samples after WGA before conducting downstream analyses. Only three samples resulted in products with less than 35% of the markers, probably due to inefficient amplification of the Phi29 polymerase-based whole-genome-amplification system, as described previously [12, 20]. 98% of the positive KASP reactions (1226 out of 1250) showed clear homozygous signals in agreement to the positive-control. 24 heterozygous calls (1.92%) were observed. These were randomly distributed across all samples and markers and therefore unlikely to be caused by an erroneous sorting of two nuclei into one microwell, so they were discarded as genotyping errors. Furthermore, no false allele calling, e.g. Barke allele instead of Morex, due to WGA errors was found in the positive controls using haploid nuclei of Morex. We conclude that multi-locus KASP-based genotyping on WGA-amplified DNA derived from single haploid nuclei is feasible.

Monitoring meiotic recombination by genotyping single pollen grains

Meiotic crossovers along chromosome 3H were measured with a mean inter-marker distance of 14.35 and 12.32 mega base-pairs (Mbp) for the short and long arm of chromosome 3H, respectively [17]. The average number of crossovers for chromosome 3H in our pollen population was 1.92, while the corresponding number for a corresponding DH population was 1.84. These two values are not significantly different (P = 0.72). Looking at the distribution of the total number of crossovers, we did not find a significantly different pattern for both pollen and DH population (Fig 3A) indicating the reliability of our approach (χ²-test, P = 0.99). Although one individual nucleus showing 6 crossovers on chromosome 3H was found to have a low genotyping call rate of 0.38, there was no significant correlation between the number of crossovers and genotyping call rate (r = -0.16, r² = 0.03, Fig 3B). We further determined recombination
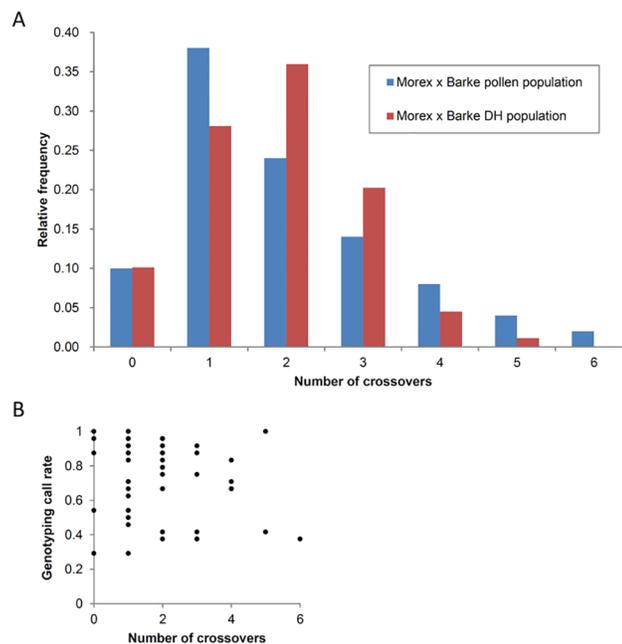
Fig 3. Comparison of the distribution of the number of crossovers. (A) The relative frequency of the total number of crossovers per chromosome 3H grouped into classes ranging from 0 to 6 of the Morex x Barke pollen population (blue) in comparison to the Morex x Barke DH population data (red) [17]. (B) Correlation between KASP genotyping call rate and the number of crossovers found for each sample (r = -0.16, r² = 0.03).

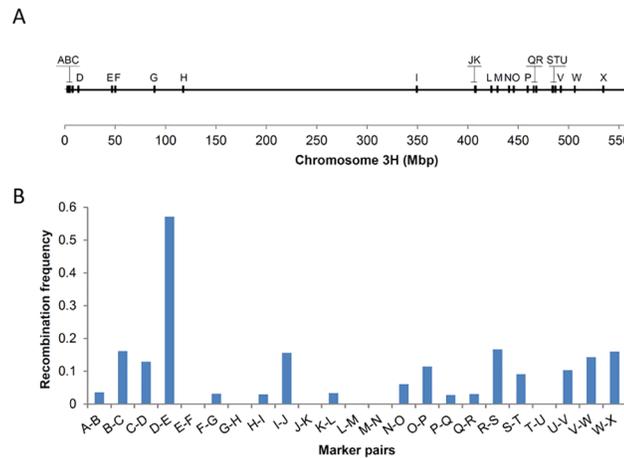doi:10.1371/journal.pone.0137677.g003

Fig 4. Recombination frequency along chromosome 3H determined by pollen genotyping. (A) KASP marker positions (A to X) are shown as vertical bars along chromosome 3H. The physical length of the chromosome (Mbp) is shown on the x-axis. **(B)** The recombination frequency along chromosome 3H of a given physical interval measured as the proportion of crossovers to no-crossovers for each marker pair. A distal bias is shown by higher recombination frequencies towards the chromosome ends and low recombination frequencies between markers H and I.

frequencies along chromosome 3H by counting the number of crossovers in neighbouring marker intervals ([Fig 4]). A typical pattern of elevated recombination frequencies towards the distal regions of the chromosome was found which is in agreement with previous reports for barley based on molecular marker data [1, 21] and cytological visualizations of crossovers [7].

To assess the extent of segregation distortion in pollen grains, we investigated the number of loci showing segregation distortion in our pollen population and compared it to equivalent data derived from a doubled haploid (DH) population of the same genotype. In our pollen population, 24 loci on chromosome 3H were scored for presence or absence of each allele. Segregation distortion was found for 8.3% (2 of 24) of the markers ([S3 Table]). This proportion appears to be lower compared to the Morex x Barke DH population which showed 25.7% (25 of 97) of all loci on chromosome 3H having distorted segregation ratios ([S3 Table]). However, the difference in sample size, which is 50 pollen grains compared to 89 DH individuals, allows only major effects to be detected. This difference might be explained by selection against particular genotype combinations during anther culture of the DH lines or by absence of selection in pollen grains for pollination and fertilization success. However, this tendency is in agreement with Sayed et al. [22] who compared segregation distortion of a barley DH population versus an F_2 population, finding a difference of 44.2% versus 16.3%, respectively.

We conclude that it is feasible to genotype single pollen grains using our amplification approach combined with KASP. It offers the opportunity to efficiently monitor meiotic recombination in individual pollen nuclei and avoids the necessity to generate segregating populations. Due to the high amount of DNA obtained from a single haploid nucleus via WGA, we suggest that our approach might be used for genome wide analyses. This will be particularly useful in plant breeding to monitor the recombination landscape of any genotype of interest.

Supporting Information

S1 Fig. Distribution of the number of positive PCR markers to assess WGA efficiency. (TIF)

S2 Fig. Distribution of KASP marker call rate and frequency of double crossover. The marker call rate of each KASP marker (A to X) is shown (blue) as well as the frequency of presumptive double crossover (red), e.g. a crossover on both sides of a given marker.
(TIF)

S1 Table. List of primers used to evaluate WGA performance.
(DOCX)

S2 Table. Gene-based KASP markers.
(DOCX)

S3 Table. Genotypic data and segregation statistics of the pollen population and DH population. Sample number is shown in the first column and the first row indicates the marker number. Genotypic values are shown as A for Morex and B for Barke. Missing values are indicated by a minus.
(XLSX)

S4 Table. Supplementary information to Figs 2, 3, 4 and S5, S6 Figs. Actual data referring to Figs 2, 3, 4 and S5, S6 Figs are summarized in separate tables.
(XLSX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: SD AH. Performed the experiments: SD JF. Analyzed the data: SD NK. Contributed reagents/materials/analysis tools: PC. Wrote the paper: SD JF AH EB.

References

1. Baker K, Bayer M, Cook N, Dreissig S, Dhillon T, Russell J, et al. The low-recombining pericentromeric region of barley restricts gene diversity and evolution but not gene expression. Plant J. 2014; 79 (6):981–92. doi: 10.1111/tpj.12600 PMID: 24947331.

2. Tondelli A, Xu X, Moragues M, Sharma R, Schnaithmann F, Ingvardsen C, et al. Structural and Temporal Variation in Genetic Diversity of European Spring Two-Row Barley Cultivars and Association Mapping of Quantitative Traits. Plant Genome-Us. 2013; 6(2). doi: 10.3835/plantgenome2013.03.0007 PMID: WOS:000322887700009.

3. Crismani W, Girard C, Froger N, Pradillo M, Santos JL, Chelysheva L, et al. FANCM limits meiotic crossovers. Science. 2012; 336(6088):1588–90. doi: 10.1126/science.1220381 PMID: 22723424.

4. Higgins JD, Perry RM, Barakate A, Ramsay L, Waugh R, Halpin C, et al. Spatiotemporal Asymmetry of the Meiotic Program Underlies the Predominantly Distal Distribution of Meiotic Crossovers in Barley. Plant Cell. 2012; 24(10):4096–109. doi: 10.1105/tpc.112.102483 PMID: ISI:000312378300018.

5. Salome PA, Bomblies K, Fitz J, Laitinen RA, Warthmann N, Yant L, et al. The recombination landscape in Arabidopsis thaliana F2 populations. Heredity. 2012; 108(4):447–55. doi: 10.1038/hdy.2011.95 PMID: 22072068; PubMed Central PMCID: PMC3313057.

6. Sybenga J. Quantitative Analysis of Chromosome Pairing and Chiasma Formation Based on Relative Frequencies of M I Configurations .4. Interchange Heterozygotes. Genetica. 1966; 37(2):199–&. doi: 10.1007/Bf01547131 PMID: WOS:A19668461900007.

7. Phillips D, Wnetrzak J, Nibau C, Barakate A, Ramsay L, Wright F, et al. Quantitative high resolution mapping of HvMLH3 foci in barley pachytene nuclei reveals a strong distal bias and weak interference.

J Exp Bot. 2013; 64(8):2139–54. doi: 10.1093/jxb/ert079 PMID: 23554258; PubMed Central PMCID: PMC3654414.

8. Copenhaver GP, Keith KC, Preuss D. Tetrad analysis in higher plants. A budding technology. Plant Physiol. 2000; 124(1):7–16. PMID: 10982416.

9. Petersen G, Johansen B, Seberg O. PCR and sequencing from a single pollen grain. Plant Mol Biol. 1996; 31:189–91. PMID: 8704154

10. Chen PH, Pan YB, Chen RK. High-throughput procedure for single pollen grain collection and polymerase chain reaction in plants. J Integr Plant Biol. 2008; 50(3):375–83. doi: 10.1111/j.1744-7909.2007.00624.x PMID: ISI:000253711300015.

11. Aziz AN, Sauve RJ. Genetic mapping of Echinacea purpurea via individual pollen DNA fingerprinting. Mol Breeding. 2008; 21(2):227–32. doi: 10.1007/s11032-007-9123-9 PMID: ISI:000251868400009.

12. Wang JB, Fan HC, Behr B, Quake SR. Genome-wide Single-Cell Analysis of Recombination Activity and De Novo Mutation Rates in Human Sperm. Cell. 2012; 150(2):402–12. doi: 10.1016/j.cell.2012.06.030 PMID: ISI:000306595700018.

13. Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science. 1983; 220(4601):1049–51. doi: 10.1126/science.220.4601.1049 PMID: 17754551.

14. De Storme N, Geelen D. The Arabidopsis mutant jason produces unreduced first division restitution male gametes through a parallel/fused spindle mechanism in meiosis II. Plant Physiol. 2011; 155 (3):1403–15. doi: 10.1104/pp.110.170415 PMID: 21257792; PubMed Central PMCID: PMC3046594.

15. Gole J, Gore A, Richards A, Chiu YJ, Fung HL, Bushman D, et al. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. Nat Biotechnol. 2013; 31 (12):1126–32. doi: 10.1038/nbt.2720 PMID: 24213699; PubMed Central PMCID: PMC3875318.

16. Comadran J, Kilian B, Russell J, Ramsay L, Stein N, Ganal M, et al. Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. Nature genetics. 2012; 44(12):1388–92. doi: 10.1038/ng.2447 PMID: 23160098.

17. International Barley Genome Sequencing Consortium (IBGSC). A physical, genetic and functional sequence assembly of the barley genome. Nature. 2012; 491(7426):711–+. doi: 10.1038/Nature11543 PMID: ISI:000311606000035.

18. Mirouze M, Lieberman-Lazarovich M, Aversano R, Bucher E, Nicolet J, Reinders J, et al. Loss of DNA methylation affects the recombination landscape in Arabidopsis. Proc Natl Acad Sci U S A. 2012; 109 (15):5880–5. doi: 10.1073/pnas.1120841109 PMID: 22451936; PubMed Central PMCID: PMC3326504.

19. Milne I, Shaw P, Stephen G, Bayer M, Cardle L, Thomas WT, et al. Flapjack—graphical genotype visualization. Bioinformatics. 2010; 26(24):3133–4. doi: 10.1093/bioinformatics/btq580 PMID: 20956241; PubMed Central PMCID: PMC2995120.

20. de Bourcy CF, De Vlaminck I, Kanbar JN, Wang J, Gawad C, Quake SR. A quantitative comparison of single-cell whole genome amplification methods. Plos One. 2014; 9(8):e105585. doi: 10.1371/journal.pone.0105585 PMID: 25136831; PubMed Central PMCID: PMC4138190.

21. Kunzel G, Korzun L, Meister A. Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics. 2000; 154(1):397–412. PMID: 10628998; PubMed Central PMCID: PMC1460903.

22. Sayed H, Kayyal H, Ramsey L, Ceccarelli S, Baum M. Segregation distortion in doubled haploid lines of barley (Hordeum vulgare L.) detected by simple sequence repeat (SSR) markers. Euphytica. 2002; 125 (2):265–72. doi: 10.1023/A:1015861610226 PMID: WOS:000176184000014.

3.2. Sequencing of single pollen nuclei reveals meiotic recombination events at megabase resolution and circumvents segregation distortion caused by postmeiotic processes

Dreissig, S. et al., 2017. Sequencing of single pollen nuclei reveals meiotic recombination events at megabase resolution and circumvents segregation distortion caused by postmeiotic processes. Front. Plant. Sci., doi: 10.3389/fpls.2017.01620

Sequencing of Single Pollen Nuclei Reveals Meiotic Recombination Events at Megabase Resolution and Circumvents Segregation Distortion Caused by Postmeiotic Processes

Steven Dreissig[1], Jörg Fuchs[1], Axel Himmelbach[2], Martin Mascher[3, 4*] and Andreas Houben[1*]

[1] Department of Breeding Research, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Germany, [2] Department of Genebank, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Germany, [3] Domestication Genomics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Germany, [4] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

Meiotic recombination is a fundamental mechanism to generate novel allelic combinations which can be harnessed by breeders to achieve crop improvement. The recombination landscape of many crop species, including the major crop barley, is characterized by a dearth of recombination in 65% of the genome. In addition, segregation distortion caused by selection on genetically linked loci is a frequent and undesirable phenomenon in double haploid populations which hampers genetic mapping and breeding. Here, we present an approach to directly investigate recombination at the DNA sequence level by combining flow-sorting of haploid pollen nuclei of barley with single-cell genome sequencing. We confirm the skewed distribution of recombination events toward distal chromosomal regions at megabase resolution and show that segregation distortion is almost absent if directly measured in pollen. Furthermore, we show a bimodal distribution of inter-crossover distances, which supports the existence of two classes of crossovers which are sensitive or less sensitive to physical interference. We conclude that single pollen nuclei sequencing is an approach capable of revealing recombination patterns in the absence of segregation distortion.

Keywords: single-cell genomics, pollen, meiosis, homologous recombination, crossover, crossover interference, segregation distortion

INTRODUCTION

Meiotic recombination is a key mechanism in eukaryotic reproduction which enables novel combinations of alleles and provides a mechanism for plant breeders to achieve crop improvement. Recombination patterns are shaped by genetic, epigenetic and environmental factors (Melamed-Bessudo and Levy, 2012; Mirouze et al., 2012; Yelina et al., 2012; Ziolkowski et al., 2015, 2017; Ritz et al., 2017). In many crops, including barley, recombination events occur predominantly in distal regions of the chromosomes where gene density is high. In contrast, interstitial and centromere-proximal regions containing 12–24% of the barley gene complement are marked by strongly reduced recombination rates (Baker et al., 2014). Although genetic diversity is reduced

in low-recombining regions, they nevertheless contain genes and thus represent a resource that is hardly accessible to plant breeders. Therefore, significant efforts are being directed toward the manipulation of recombination frequency and distribution. Several approaches were shown to be successful, including the increase of crossovers via mutation of an anti-crossover factor (Crismani et al., 2012), epigenetic remodeling of crossover frequency via reduced DNA methylation (Melamed-Bessudo and Levy, 2012; Mirouze et al., 2012; Yelina et al., 2012; Habu et al., 2015), and shifting of crossover positions via increased or decreased temperatures (Higgins et al., 2012; Phillips et al., 2015; Martin et al., 2017). Furthermore, natural diversity of recombination patterns was shown to exist in *Arabidopsis*, maize, and *Hordeum* (Gale et al., 1970; Sall, 1990; Sall et al., 1990; Nilsson and Pelger, 1991; Sidhu et al., 2015; Ziolkowski et al., 2015, 2017).

In addition to low recombining regions limiting crop improvement, segregation distortion (SD) is another undesirable phenomenon as it reduces the chance of combining certain alleles. SD is defined as a deviation of the segregation ratio of alleles from the expected Mendelian segregation ratio. In barley double haploid (DH) populations, large proportions of the genome can show segregation distortion (Bélanger et al., 2016a). A frequent cause of segregation distortion is selection acting on genetically linked loci which results in entire chromosomal regions showing segregation distortion (hereafter termed SDR for segregation distortion region) (Hiraizumi et al., 1960; Hill and Robertson, 1966).

Taken together, tight genetic linkage of large proportions of the genome and distorted segregation resulting in a linkage drag of alleles hamper the advance of plant breeding. Future attempts to overcome these restrictions will require efficient methods to assay such effects. There are numerous methods to measure meiotic recombination in plants, including molecular markers (Salome et al., 2012), cytological visualization of crossovers (Sybenga, 1966; Anderson et al., 2003; Phillips et al., 2013), tetrad analysis (Copenhaver et al., 2000), fluorescent protein-tagged loci expressed in pollen (Yelina et al., 2013), and several pollen genotyping approaches (Drouaud and Mezard, 2011; Khademian et al., 2013; Dreissig et al., 2015). Although these methods have been successfully used to characterize recombination patterns and improve our understanding of meiosis, each of them has its specific advantages and disadvantages. The analysis of recombination by molecular markers requires the generation of a segregating population, which is laborious and very challenging for some plant species. Cytological analysis of recombination is more widespread and applicable to many plant species, yet its resolution is lower compared to sequence-based approaches and the analysis is demanding in terms of time and experience. Tetrad analysis combined with fluorescence markers is a very powerful high-throughput approach but requires the integration of reporter transgenes and is so far limited to the model species *Arabidopsis*.

Single-cell sequencing is a new technology that holds the promise to directly measure the outcome of meiosis in individual cells, e.g., microspores (Li et al., 2015) or pollen grains. We have previously developed a single pollen genotyping approach based on flow-sorting of haploid nuclei followed by whole genome

amplification via multiple-displacement-amplification (MDA) of DNA and multi-locus competitive allele specific PCR (KASP) genotyping (Dreissig et al., 2015). This approach has shown the potential of single-cell analyses to measure recombination, but was limited by the number of KASP markers that could be assayed. To overcome this restriction, we took advantage of representative whole-genome amplification combined with next-generation-sequencing (NGS) library preparation and sequencing in the current study.

Here we present a new approach to directly investigate meiotic recombination at the DNA sequence level by combining flow-sorting of pollen nuclei with PicoPLEX single-cell sequencing (Rubicon Genomics). This sequencing approach is based on quasi-random PCR amplification of single-cell genomic DNA and yields a library with dual indexes for limited coverage sequencing. We show that this approach is capable of measuring meiotic recombination and segregation ratios throughout the whole genome of the large genome species barley at megabase resolution by comparing our results obtained through pollen sequencing to genotyping-by-sequencing (GBS) data of a barley DH population.

MATERIALS AND METHODS

Plant Material and Isolation of Single Pollen Nuclei

Pollen grains were collected from a *Hordeum vulgare* L. F_1 plant derived from a cross between the cultivars "Morex" ($\male$) and "Barke" ($\female$) and grown at 20°C during the day (7:00–20:00) and 16°C during the night. Pollen nuclei were isolated and stained as described previously (Dreissig et al., 2015) and sorted using a BD Influx cell sorter (BD Biosciences) into a 384 microwell plate (Applied Biosystems) using the "1.0 drop single" sort mode of the BD FACS software. As a control, we sorted three individual pollen nuclei from the parental genotype "Barke."

Single Nuclei Library Preparation and Illumina Sequencing

Illumina NGS libraries were prepared from 43 individual nuclei using the PicoPLEX DNA-seq kit essentially following the manufacturer's instructions (Rubicon Genomics). After the final amplification reaction with primers containing unique dual barcodes suitable for Illumina NGS, 10 μl aliquots of each library were pooled. The pooled DNA sample was purified using AMPure XP beads (Beckman Coulter Inc.) as described (Rubicon Genomics). The pool was eluted in 30 μl TE (pH 8.0) and size-fractionated using a SYBR-Gold stained 2% agarose gel (Himmelbach et al., 2014). The region of interest (350–1,000 bp) was excised, and the DNA was extracted using the Qiagen MinElute Kit (Himmelbach et al., 2014). The library was characterized using an Agilent 2100 Bioanalyzer (Himmelbach et al., 2014) and quantified by Real-Time PCR as described (Mascher et al., 2013b). After the addition of 8% PhiX DNA as a control, the pooled library was sequenced using the Illumina HiSeq2500 device (rapid run, 1 lane, cBot clustering, 2x 100 cycles paired-end, dual-indexing with 8 cycles per index) according to

the manufacturer's instructions. Sequence raw data are available under EMBL ENA accession PRJEB21630.

Sequence Read Mapping and Genotype Analysis

Illumina adapters were trimmed using Cutadapt version 1.12 (Martin, 2011). Trimmed reads were aligned to the barley cv. "Morex" reference genome sequence assembly (Mascher et al., 2017) using BWA-MEM version 0.7.15 (Li, 2013) with default parameters. The resulting SAM files were converted to BAM format with SAMtools (Li et al., 2009). Sorting and detection of optical and PCR duplicates was done with Novosort (http://www.novocraft.com/products/novosort/). SAMtools version 1.3 (Li, 2011) was used for multiple-sample genotype calling at single-nucleotide polymorphism (SNP) sites which were previously ascertained in the "Morex" × "Barke" RIL population using the POPSEQ method (Mascher et al., 2013a). VCF files were imported into the R statistical environment (R Core Team, https://www.r-project.org/contributors.html). Consensus genotypes were derived by aggregating information in 1 Mb bins using functionalities of the R package "data.table" (https://cran.r-project.org/package=data.table). This resulted in a genotype file containing allele information at 1 megabase pair (Mbp) resolution which was used to analyse recombination frequency and segregation distortion.

We used GBS data derived from a "Morex" × "Barke" DH population which was described previously (IBGSC, 2012) for comparison. GBS data were retrieved from https://wheat.pw.usda.gov/ggpages/MxB/. GBS tags were mapped onto the most recent version of the barley reference genome sequence (Mascher et al., 2017) an aggregated in 1 Mbp intervals.

Recombination Analysis Based on Pollen and a Double Haploid Population

To identify meiotic recombination events in the pollen and double haploid (DH) population, we searched for recombination patterns in each genotype matrix which were indicated by changes from "0" ("Barke" allele) to "2" ("Morex" allele) or vice versa. To count recombination events, we conducted a text search for patterns indicating recombination events (e.g., $0{\to}0{\to}0{\to}2{\to}2{\to}2$). We manually curated the genotype files by removing markers showing a high frequency of double crossovers (e.g., $0{\to}2{\to}0$), which were considered genotyping errors (Salome et al., 2012). To map the approximate position of recombination events onto the physical map of the barley genome, a 5-Mbp sliding window approach was used to scan along each chromosome searching for allele changes from "0" to "2" and vice versa. We then calculated recombination frequency in cM/Mbp [cM = 100*(# of recombinations/#total)] along each chromosome by counting the number of recombination events in 5-Mbp sliding windows relative to the total number of samples. To analyse crossover interference, we extracted all samples showing more than two recombination events on a given chromosome and calculated the physical distance (Mbp) between nearby recombination events. To determine the effect of crossover interference, we used the crossover

distribution analyser (CODA) software (Gauthier et al., 2011) which compares observed inter-crossover distances against a simulated gamma model to calculate nu. A value of $nu = 1$ indicates no interference, $nu < 1$ indicates negative interference, and $nu > 1$ indicates positive interference. Genotype data are available as Supplementary File 1.

Analysis of Segregation Distortion in Pollen and Double Haploid Population

Segregation distortion was analyzed by calculating average allele frequencies in 10 Mbp sliding windows along each chromosome of both populations. Markers with >50% missing data were removed from the analysis. To test for significant deviation from the expected segregation ratio of 1:1 of each parental allele, we conducted a χ^2-test between expected and observed allele frequencies. Segregation distortion regions (SDR) were identified by a significant deviation from the expected ratio of 1:1 ($P < 0.05$).

RESULTS

Sequencing of Individual Pollen Nuclei

To identify recombination events, we first sequenced the genomes of individual haploid pollen nuclei. Toward this purpose, we utilized our previously established approach for pollen nuclei isolation (Dreissig et al., 2015) combined with PicoPLEX single-cell DNA amplification and NGS library preparation. A total of 40 pollen nuclei derived from a single "Morex" (♂) x "Barke" (♀) F_1 plant were subjected to PicoPLEX sequencing. As a control, pollen nuclei obtained from the parental genotype "Barke" were used. The initial DNA amplification via quasi-random priming yielded an average fragment size of 933 bp. No amplification was detected in the negative control which indicates that the amount of DNA contamination was below the level of detection. Sequencing the 40 pollen nuclei on the Illumina HiSeq 2500 platform yielded between 2.7 million and 11.6 million (mean: 5.9 million) reads per sample, corresponding to an average read depth of 0.1x per haploid nucleus. Reads were mapped to the reference genome assembly of cv. "Morex" (Mascher et al., 2017) and genotypes were called at single-nucleotide polymorphism (SNP) sites known to segregate in the "Morex" × "Barke" population (Mascher et al., 2013a). Consensus genotypes were derived by aggregating SNP information in 1 Mbp bins based on the reference genome. **Figure 1** shows the graphical genotypes of the 40 pollen nuclei at 1 Mbp resolution.

Comparing the Recombination Landscape of Barley Pollen and DH Plants

Based on cytological analyses (Sybenga, 1966; Phillips et al., 2013; Aliyeva-Schnorr et al., 2015) and molecular analyses of segregating populations (Künzel et al., 2000; IBGSC, 2012; Phillips et al., 2015), the recombination landscape of barley is characterized by elevated recombination frequencies in distal chromosome regions and strongly reduced recombination in (peri-)centromeric regions. In order to overcome the resolution limit of cytological analyses, we attempted to investigate the
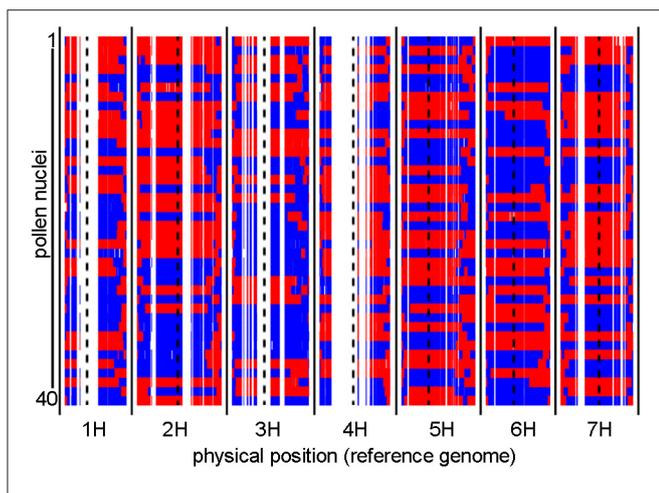
FIGURE 1 | Graphical genotypes of individual pollen nuclei revealed by single-cell genome sequencing. Recombination events were detected in 40 individual pollen nuclei. The two parental barley genotypes are shown in red ("Morex") and blue ("Barke"). Consensus genotypes were mapped to the physical reference genome of barley at 1 Mbp resolution. Centromere positions are indicated by dashed black lines. White gaps which consistently occur in all samples are regions where no genetic polymorphisms exist between "Morex" and "Barke."
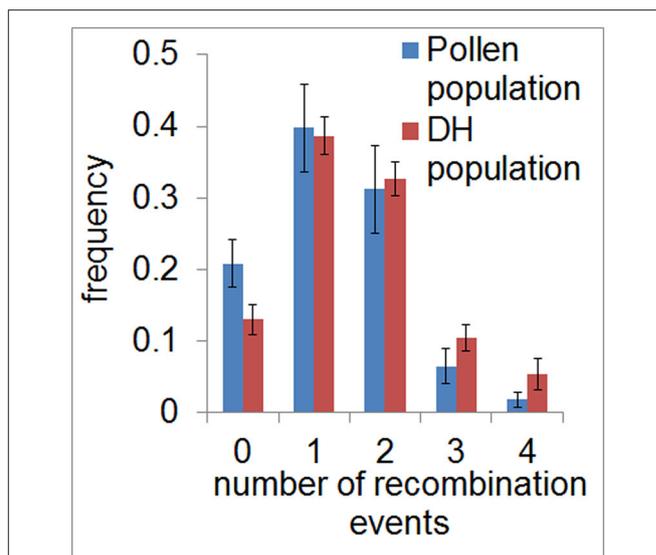

FIGURE 2 | Frequency of recombination events in pollen and DH plants. Relative frequency of the average number of recombination events per chromosome is shown for the pollen (blue) and DH population (red) in classes ranging from 0 to 4. Error bars represent the standard deviations based on measurements conducted on all seven barley chromosomes.

FIGURE 3 | Elevated recombination frequencies in distal regions of barley chromosome 5H. Recombination frequency in pollen (blue) and DH plants (red) was calculated in 5 Mbp sliding windows along chromosome 5H and plotted along the physical map. The position of the centromere is marked by a black diamond.

recombination landscape of barley directly at the DNA sequence level by sequencing individual pollen nuclei.

To assess the recombination landscape of barley pollen compared to DH plants, we first counted the number of recombination events in each sample in both populations. We measured a total of 380 recombination events in the population of 40 haploid pollen nuclei (average of 9.5 per pollen nucleus, SE = 0.38) and 974 recombination events in the DH population composed of 89 plants (average of 10.9 per DH plant, SE = 0.3). Predominantly, we detected one or two recombination events per chromosome in both populations with 38.7–39.8% of samples showing one recombination event and 31.1–32.6% of samples showing two recombination events. The number of recombination events, which was ranging from zero to four per chromosome, was found to be similar between pollen and DH population (χ^2-goodness of fit test, $P > 0.99978$) (**Figure 2**). The occurrence of chromatids apparently lacking any recombination event detected by SNPs (13–20%) seems to be the same as in an *Arabidopsis* data set described by Salome et al. (2012). Consequently, recombination frequency was found to be similar in barley pollen compared to whole DH plants.

Since the number of recombination events per chromosome was highly similar between the pollen population and the DH population, we then examined whether the genome wide distribution of recombination events differed between both populations. We measured recombination frequencies along all chromosomes of barley using a 5 Mbp sliding window approach. In both populations, we found elevated recombination frequencies in distal regions of all chromosomes and almost no recombination in (peri-)centromeric regions (**Figure 3**, Supplementary files 2–7). This observation is in agreement with
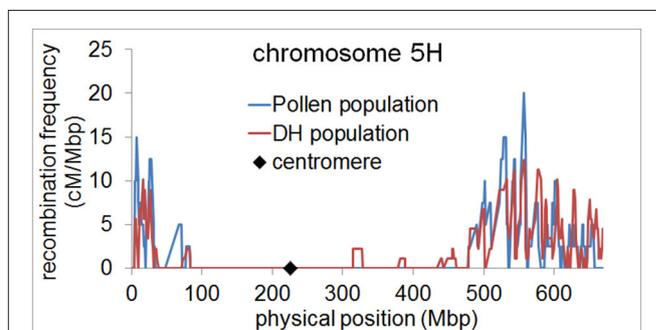
previous studies showing a skewed distribution of recombination events toward distal chromosome regions in barley (Künzel, 1982; Linde-Laursen, 1982; Künzel et al., 2000; Phillips et al., 2013; Baker et al., 2014; Dreissig et al., 2015). It also shows that there is no different positioning of recombination events in pollen, i.e., in (peri-)centromeric regions. These regions were shown to harbor essential genes encoding proteins for basic cellular functions such as translation and photosynthesis (Mascher et al., 2017). It could therefore be reasoned that (peri-)centromeric recombination events could theoretically be absent in DH plants due to selection against housekeeping gene-encoding (peri-)centromeric sites of recombination which would disrupt linkage between essential genes.

In agreement with the predominantly distal positioning of recombination events in both populations, we found positive crossover interference indicated by 48.9–59.8% of recombination events being separated by more than 400 Mbp (range = 402–729 Mbp) over a chromosome size ranging from 558 to 767 Mbp. Interestingly, 35.6–39.6% of recombination events were separated by less than 100 Mbp (range = 10–98 Mbp) (**Figure 4**). The smallest distance between two recombination events was 10 Mbp which corresponds to ~1.5% of the chromosome. We conducted a crossover interference analysis (gamma model; measured in nu) to determine the strength of interference (Gauthier et al., 2011). A value of $nu = 1$ indicates no interference, $nu < 1$ indicates negative interference, and $nu > 1$ indicates positive interference. Due to the low number of chromosomes showing at least two recombination events, we did not analyse chromosomes separately, but pooled data from all seven barley chromosomes. Positive interference values of $nu = 4.76$ and 3.02 were detected in DH and pollen populations, respectively. In addition, we split all recombination events into two groups with <100 or >400 Mbp distance between two events. When both groups were analyzed separately, we found weaker interference values for recombination events less than 100 Mbp apart ($nu = 2.336$ for pollen and $nu = 2.202$ for DH population) and stronger interference values when more than 400 Mbp apart ($nu = 8.511$ for pollen and $nu = 8.199$ for DH population). These patterns might be attributed to interference sensitive and less sensitive crossovers, i.e., class I and class II crossover. We then tested whether recombination events separated by less than 100 Mbp were confined to specific chromosomal regions or distributed randomly by plotting the physical positions of multiple recombination events on the same chromosome against themselves (**Figure 5**). All recombination events separated by less than 100 Mbp were strictly confined to distal regions, which corresponds to the accumulation of dots in the bottom left and top right quarters of **Figure 5**. Recombination events separated by more than 400 Mbp were located on different arms (dots in the top left quarter of **Figure 5**). Our data show that crossover interference is positive in barley. However, a substantial proportion of recombination events is separated by less than 100 Mbp which supports the existence of class I and class II crossovers in barley.

Segregation Distortion Is High in DH Plants, But Almost Absent in Pollen

Segregation distortion is defined as the preferential transmission of one allele over the other, which results in a statistically significant deviation from an expected Mendelian segregation ratio of 1:1. We asked whether the extent of segregation distortion differs between pollen and DH plants. Our hypothesis was that segregation distortion would be substantially lower in pollen because of the absence of any selective pressure which might arise during pollen tube growth, fertilization, hybrid compatibility, and plant development. We expected the opposite in the DH population because of selective pressure during microspore culture, embryo development, plant regeneration, and spontaneous diploidization. It is important to note that the
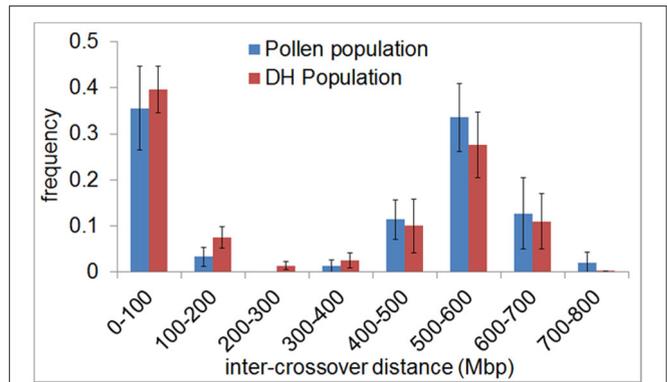
FIGURE 4 | Inter-crossover distance reveals positive crossover interference and supports the existence of two crossover classes in barley. The frequency of the distance between crossovers on the same chromatid (inter-crossover distance) in pollen (blue) and DH plants (red) was determined in 100 Mbp classes ranging from <100 to >700 Mbp. The relative frequency of nearby crossovers present in each class was plotted. Error bars represent the standard deviation based on measurements conducted on all seven barley chromosomes.
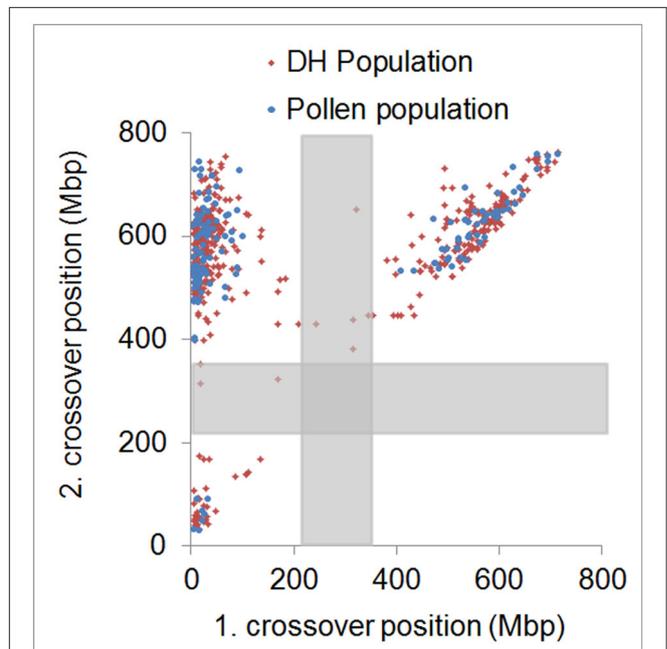
FIGURE 5 | Physical distribution of first and second crossover positions. Physical positions of first and second crossover event for all samples showing more than two crossovers in the pollen (blue) and DH (red) population. Approximate centromeric regions are marked by gray boxes. Strong physical interference is shown by dots accumulated in the top left quarter. Weak physical interference is shown by dots accumulated in the bottom left and top right quarter.

DH population which was genotyped and provided by the IBGSC (2012) consisted of spontaneously diploidized plants only.

In the pollen population, we found normal segregation ratios for almost all chromosomal regions (Supplementary files 8–12). The exceptions were one region on chromosome 2H located at

736–752 Mbp and two regions on chromosome 3H located at 634–642 Mbp and 682–695 Mbp (**Figure 6**). These regions only amount to 2 and 3% of chromosome 2H and 3H, respectively. In both cases, these SDRs were located in high recombining regions of the chromosome allowing them to remain small and not cause distorted segregation of a larger part of the chromosome through linkage (Supplementary file 13). In contrast, in the DH population, a high proportion of large chromosomal regions were affected by segregation distortion. We detected a total of 15 SDRs distributed across all chromosomes which varied in size ranging from 0.01 up to 87.3% of the chromosome. Major SDRs, varying from 72.6 up to 87.3% of the chromosome, were found on chromosome 1H, 2H, 5H, and 7H (**Figure 6A**, Supplementary files 8, 10, 12). In addition to these major SDRs, we detected 11 minor SDRs which varied in size ranging from 0.01 up to 5% of the chromosome (**Figure 6B**, Supplementary files 8, 10–12). Interestingly, we did not detect the same SDRs on chromosome 2H and 3H in the pollen population as in the DH population which indicates different selective pressures acting on these loci. For example, in the DH population, two regions of chromosome 3H (571.6–606.6 Mbp and 672.2–698.3 Mbp) exhibited higher transmission of the "Morex" allele whereas, in the pollen population, two regions of the chromosome (634–642 Mbp and 682–695 Mbp) exhibited higher transmission of the "Barke" allele (**Figure 6B**). This example shows that under varying conditions (e.g., pollen development vs. DH production) not only different regions can be selected, but also different parental alleles can be preferentially transmitted.

Hence, our results show that segregation distortion is almost absent in pollen grains which supports the conclusion that meiosis alone is not the main cause of this phenomenon. On the contrary, segregation distortion was found for nearly half of the entire genome (49.9%) in barley DH plants. We conclude that selective pressure during microspore culture, embryo development, plant regeneration, and diploidization is the most likely cause for segregation distortion in DH plants.

DISCUSSION

The main conclusion of the present study is that the recombination landscape of barley pollen and DH plants does not differ in frequency or positioning of recombination events, yet segregation distortion is almost absent in pollen grains whereas it is detectable to a large extent in DH plants likely caused by selection during DH production. In addition, we present recombination measurements which support the existence of class I and class II crossovers in barley. We demonstrate that our approach for single pollen nuclei sequencing is suitable to directly investigate the recombination landscape of barley at the molecular level in an unbiased way.

Pollen Sequencing as a Robust Approach to Directly Measure Recombination at Megabase Resolution in Barley

We sought to analyse recombination in pollen and DH plants separately to test if the typical recombination pattern
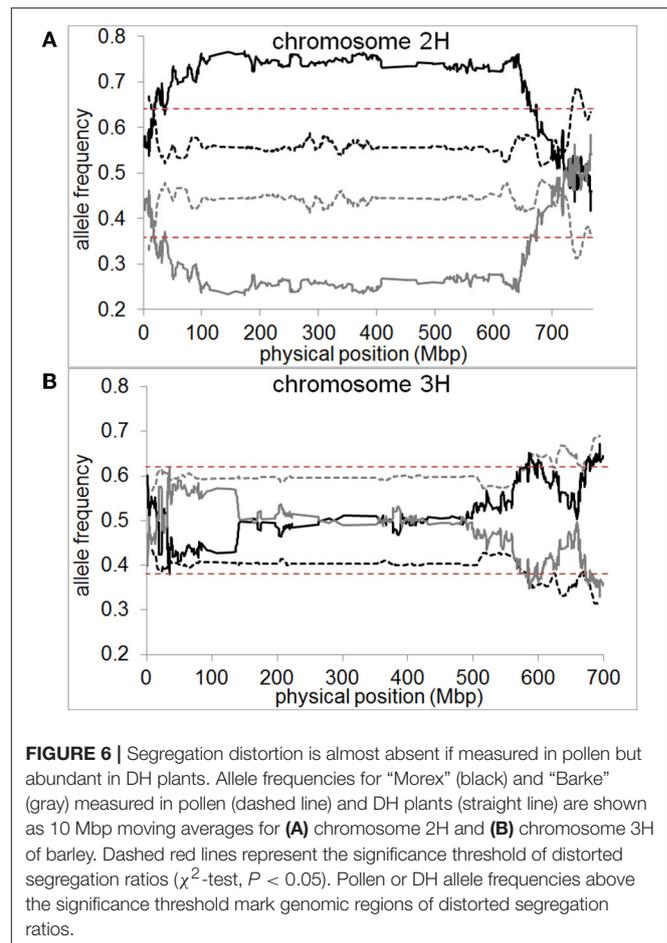
FIGURE 6 | Segregation distortion is almost absent if measured in pollen but abundant in DH plants. Allele frequencies for "Morex" (black) and "Barke" (gray) measured in pollen (dashed line) and DH plants (straight line) are shown as 10 Mbp moving averages for **(A)** chromosome 2H and **(B)** chromosome 3H of barley. Dashed red lines represent the significance threshold of distorted segregation ratios (χ^2-test, $P < 0.05$). Pollen or DH allele frequencies above the significance threshold mark genomic regions of distorted segregation ratios.

found in segregating populations of barley, characterized by a predominantly distal positioning of recombination events, is caused by selection against (peri-)centromeric recombination events or reflects the real outcome of meiosis. The low recombining regions of the barley genome were previously shown to constrain gene diversity (IBGSC, 2012; Baker et al., 2014). This phenomenon is widespread in nature and is most likely caused by a combination of selective sweeps via fixation of advantageous alleles and background selection against deleterious mutations (Hill and Robertson, 1966; Smith and Haigh, 1974; Hudson, 1994; Wright et al., 2006). Furthermore, it was recently shown that essential genes involved in translation and photosynthesis reside in (peri-)centromeric low-recombining regions of the barley genome (Mascher et al., 2017). It could thus be argued that recombination events in low-recombining regions would break linkage between advantageous alleles and therefore be selected against. In pollen, however, these recombination events could still be present due to the absence of selective pressure which certainly arises during pollen tube growth, fertilization, and plant development (Pedersen, 1988; Sarigorla et al., 1992; Walsh and Charlesworth, 1992).

Our data show that the recombination landscape of barley, characterized by elevated recombination frequencies in distal regions (**Figure 3**), is truly the outcome of meiosis and not

a result of postmeiotic selection against (peri-)centromeric recombination events. This is in agreement with previous cytogenetic studies taking direct recombination measurements by means of scoring MHL3 immunostaining foci or chiasmata (Bennett et al., 1973; Phillips et al., 2013). However, it was of interest for us to test if these observations reveal the same recombination landscape as by sequencing of pollen nuclei. The direct sequencing of pollen nuclei, through the approach presented in this study, offers a much higher resolution in detecting the positions of recombination events (i.e., 1 Mbp, approximately 0.2% of the smallest barley chromosome) compared to the mapping of MLH3 fluorescence foci during meiotic prophase by structured illumination microscopy (Phillips et al., 2013). Compared to chiasmata counts performed in a variety of barley genotypes, the average number of recombination events detected in our study seems to be lower (Gale et al., 1970; Bennett et al., 1973; Colas et al., 2016). If it holds true that all cytologically defined chiasmata represent genetic exchanges between homologous chromosomes, we cannot exclude that certain recombination events are missing in our data sets. On the other hand, we measured similar recombination frequencies in pollen and DH plants while both populations were genotyped by two different methods, i.e., single-cell sequencing vs. genotyping-by-sequencing of DH plants. Furthermore, both approaches are based on haploid male gametes where only one of the four possible meiotic products, i.e., chromatids, is present. Hence, as evident from *Arabidopsis* tetrad analysis where all four chromatids are analyzed (Lu et al., 2012; Wijnker et al., 2013), it is possible for a haploid pollen nucleus to contain the exact chromatid that did not undergo meiotic recombination. It is therefore unlikely that single cell sequencing accounts for missing recombination events. It could also be argued that these differences reflect genotypic variations or environmental effects as such were shown in many cases (Sall et al., 1990; Bauer et al., 2013; Phillips et al., 2015; Sidhu et al., 2015; Ziolkowski et al., 2015, 2017).

We detected positive crossover interference in both pollen and DH plants, which is in agreement with the primarily distal positioning of recombination events. Previously, Phillips et al. (2013) reported for barley that 34–38% of crossovers are <20% of chromosome length apart and the majority of crossovers are >70% apart which results in a bimodal distribution of inter-crossover distances. Here, we found 36.8–40.4% of crossovers separated by less than 100 Mbp (approximately 15% of chromosome length) and 48.3–57.4% separated by more than 400 Mbp (approximately 60% of chromosome length) reflecting a similar bimodal distribution of inter-crossover distances (**Figure 4**). The minimum inter-crossover distance found in our study was 10 Mbp which refers to 1.5% of the corresponding chromosome. We quantified crossover interference strength (gamma model; measured in *nu*) in the pollen and DH population. We detected positive physical interference between crossovers in both pollen (*nu* = 3.02) and DH population (*nu* =4.76). These interference values are higher than those previously reported for the barley cultivar "Morex," which was at *nu* = 1.58 (Phillips et al., 2013). However, Higgins et al. (2014) argued that crossover interference might actually be stronger than estimated by Phillips et al. (2013) because the

relative separation of MLH3 foci was measured when synapsis of chromosomes was completed and not at the exact time point when crossover designation took place during synapsis. Our data, which are based on scoring crossovers at the sequence level, support this hypothesis by showing stronger crossover interference values for barley.

The existence of two crossover classes, namely class I for interference-sensitive crossovers and class II for interference-insensitive crossovers, was shown in *S. cerevisiae* and *A. thaliana* mutants being defective for core components involved in class I crossover formation (Börner et al., 2004; Higgins et al., 2004). In these mutants, 15% of crossovers of the wild-type level were still formed, which indicates the existence of an alternative class II pathway. However, the presence of two crossover classes has not been confirmed experimentally in barley yet although increasing evidence supports their existence (Phillips et al., 2013, 2015). In our study, the occurrence of recombination events separated by <100 or >400 Mbp supports the existence of interference-sensitive and less sensitive crossovers, i.e., class I and class II. However, it remains a matter of speculation why nearby crossovers are strictly confined to distal regions and do not span (peri-)centromeric regions. There is a well-known correlation between low-recombining (peri-)centromeric regions and certain histone modifications in barley, i.e., histone H3K9me2, H3K9me3, H3K27me1, and H3K27me2, as shown by chromatin immunoprecipitation (ChIP) sequencing in barley seedlings (Baker et al., 2015). Furthermore, it was shown in *Arabidopsis* that DNA methylation restricts crossovers in centromeric regions and that crossover hot spots are associated with active chromatin modifications such as H2A.Z and H3K4me3 (Yelina et al., 2012; Choi et al., 2013). It could therefore be argued that by changing specific DNA or histone modifications, crossover positioning could be manipulated to increase genetic recombination in (peri-)centromeric regions in crops such as barley.

Comparison of Segregation Distortion in Pollen and DH Plants

Segregation distortion is a widespread phenomenon in plant populations characterized by a deviation from the expected Mendelian segregation ratio. For plant breeders, it presents a problem as it has an effect on allele frequencies and can reduce the chances of obtaining specific combinations of alleles. Double haploid technology has developed into one of the most important methods for plant breeders to accelerate the otherwise lengthy process of obtaining homozygous genotypes (Germana, 2011). The disadvantage of this technology is that it is accompanied by segregation distortion to a very high extent in many genotypes and species (Xu et al., 1997; Taylor and Ingvarsson, 2003; Bélanger et al., 2016a). Segregation distortion during DH production appears to be caused by selective pressure acting upon certain loci or genomic regions. Selective pressure might arise during microspore culture, embryogenesis, plant regeneration, and spontaneous diploidization of haploid plants. Bélanger et al. (2016b) have shown that segregation distortion in barley arises predominantly during embryogenesis and plant regeneration.

In the current study, we hypothesized that segregation distortion would be low if measured in pollen grains due to the absence of selective pressure. Our data show that only three small chromosomal regions show distorted segregation ratios in pollen, amounting to 0.8% of the genome, whereas nearly 50% of the genome shows distorted segregation ratios in DH plants. This suggests that segregation distortion is not a direct outcome of meiosis but a product of selection acting at different developmental stages. Compared to Bélanger et al. (2016b) who detected no segregation distortion in immature pollen, we found one region on chromosome 2H and two regions on chromosome 3H with distorted segregation rations in mature pollen. It can be speculated that these regions might play a role in pollen development and therefore show distorted segregation. Furthermore, environmental conditions, e.g., heat stress (Frova and Sari-Gorla, 1994) or higher nutrient levels in the soil (Martin et al., 2017) can have an effect on segregation ratios in pollen, although our experiment did not involve any stress treatment.

Further improvements in protocols and decreases in the price of sequencing should enable the application of single pollen sequencing as a novel prediction tool in research and plant breeding in a wide range of species.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2017. 01620/full#supplementary-material

REFERENCES

Aliyeva-Schnorr, L., Beier, S., Karafiatova, M., Schmutzer, T., Scholz, U., Dolezel, J., et al. (2015). Cytogenetic mapping with centromeric bacterial artificial chromosomes contigs shows that this recombination-poor region comprises more than half of barley chromosome 3H. *Plant J.* 84, 385–394. doi: 10.1111/tpj.13006

Anderson, L. K., Doyle, G. G., Brigham, B., Carter, J., Hooker, K. D., Lai, A., et al. (2003). High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. *Genetics* 165, 849–865.

Baker, K., Bayer, M., Cook, N., Dreissig, S., Dhillon, T., Russell, J., et al. (2014). The low-recombining pericentromeric region of barley restricts gene diversity and evolution but not gene expression. *Plant J.* 79, 981–992. doi: 10.1111/tpj.12600

Baker, K., Dhillon, T., Colas, I., Cook, N., Milne, I., Milne, L., et al. (2015). Chromatin state analysis of the barley epigenome reveals a higher-order structure defined by H3K27me1 and H3K27me3 abundance. *Plant J.* 84, 111–124. doi: 10.1111/tpj.12963

Bauer, E., Falque, M., Walter, H., Bauland, C., Camisan, C., Campo, L., et al. (2013). Intraspecific variation of recombination rate in maize. *Genome Biol.* 14:R103. doi: 10.1186/gb-2013-14-9-r103

Bélanger, S., Clermont, I., Esteves, P., and Belzile, F. (2016a). Extent and overlap of segregation distortion regions in 12 barley crosses determined via a pool-GBS approach. *Theor. Appl. Genet.* 129, 1393–1404. doi: 10.1007/s00122-016-2711-5

Bélanger, S., Esteves, P., Clermont, I., Jean, M., and Belzile, F. (2016b). Genotyping-by-sequencing on pooled samples and its use in measuring segregation bias during the course of androgenesis in barley. *Plant Genome* 9. doi: 10.3835/plantgenome2014.10.0073

Bennett, M. D., Finch, R. A., Smith, J. B., and Rao, M. K. (1973). Time and duration of female meiosis in wheat, rye and barley. *Proceed. Royal Soc. B Biol. Sci.* 183, 301–319. doi: 10.1098/rspb.1973.0019

Börner, G. V., Kleckner, N., and Hunter, N. (2004). Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis. *Cell* 117, 29–45. doi: 10.1016/S0092-8674(04)00292-2

Choi, K., Zhao, X., Kelly, K. A., Venn, O., Higgins, J. D., Yelina, N. E., et al. (2013). *Arabidopsis* meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nat. Genet.* 45, 1327–1336. doi: 10.1038/ng.2766

Colas, I., Macaulay, M., Higgins, J. D., Phillips, D., Barakate, A., Posch, M., et al. (2016). A spontaneous mutation in MutL-Homolog 3 (HvMLH3) affects synapsis and crossover resolution in the barley desynaptic mutant des10. *New Phytol.* 212, 693–707. doi: 10.1111/nph.14061

Copenhaver, G. P., Keith, K. C., and Preuss, D. (2000). Tetrad analysis in higher plants. A budding technology. *Plant Physiol.* 124, 7–16. doi: 10.1104/pp.124.1.7

Crismani, W., Girard, C., Froger, N., Pradillo, M., Santos, J. L., Chelysheva, L., et al. (2012). FANCM limits meiotic crossovers. *Science* 336, 1588–1590. doi: 10.1126/science.1220381

Dreissig, S., Fuchs, J., Capal, P., Kettles, N., Byrne, E., and Houben, A. (2015). Measuring meiotic crossovers via multi-locus genotyping of single pollen grains in barley. *PLoS ONE* 10:e0137677. doi: 10.1371/journal.pone. 0137677

Drouaud, J., and Mezard, C. (2011). Characterization of meiotic crossovers in pollen from *Arabidopsis thaliana*. *Methods Mol. Biol.* 745, 223–249. doi: 10.1007/978-1-61779-129-1_14

Frova, C., and Sari-Gorla, M. (1994). Quantitative trait loci (QTLs) for pollen thermotolerance detected in maize. *Mol. Gen. Genet.* 245, 424–430. doi: 10.1007/BF00302254

Gale, M. D., Rees, H., and Others (1970). Genes controlling chiasma frequency in *Hordeum*. *Heredity (Edinb)*. 25, 393–410. doi: 10.1038/hdy.1970.40

Gauthier, F., Martin, O. C., and Falque, M. (2011). CODA (crossover distribution analyzer): quantitative characterization of crossover position patterns along chromosomes. *BMC Bioinform.* 12:27. doi: 10.1186/1471-2105-12-27

Germana, M. A. (2011). Gametic embryogenesis and haploid technology as valuable support to plant breeding. *Plant Cell Rep.* 30, 839–857. doi: 10.1007/s00299-011-1061-7

Habu, Y., Ando, T., Ito, S., Nagaki, K., Kishimoto, N., Taguchi-Shiobara, F., et al. (2015). Epigenomic modification in rice controls meiotic recombination and segregation distortion. *Mol. Breed.* 35:103. doi: 10.1007/s11032-015-0299-0

Higgins, J. D., Armstrong, S. J., Franklin, F. C., and Jones, G. H. (2004). The *Arabidopsis* MutS homolog AtMSH4 functions at an early step in recombination: evidence for two classes of recombination in *Arabidopsis*. *Genes. Dev.* 18, 2557–2570. doi: 10.1101/gad.317504

Higgins, J. D., Osman, K., Jones, G. H., and Franklin, F. C. (2014). Factors underlying restricted crossover localization in barley meiosis. *Annu. Rev. Genet.* 48, 29–47. doi: 10.1146/annurev-genet-120213-092509

Higgins, J. D., Perry, R. M., Barakate, A., Ramsay, L., Waugh, R., Halpin, C., et al. (2012). Spatiotemporal asymmetry of the meiotic program underlies the predominantly distal distribution of meiotic crossovers in Barley. *Plant Cell* 24, 4096–4109. doi: 10.1105/tpc.112.102483

Hill, W. G., and Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genet. Res.* 8, 269–294. doi: 10.1017/S0016672300010156

Himmelbach, A., Knauft, M., and Stein, N. (2014). Plant sequence capture optimised for Illumina sequencing. *Bio. Protocol.* 4:e1166. doi: 10.21769/BioProtoc.1166

Hiraizumi, Y., Sandler, L., and Crow, J. E. (1960). Meiotic drive in natural populations of *Drosophila melanogaster*.3. Populational implications of the segregation-distorter locus. *Evolution* 14, 433–444.

Hudson, R. R. (1994). How can the low levels of DNA sequence variation in regions of the *Drosophila* genome with low recombination rates be explained? *Proc. Natl. Acad. Sci. U.S.A* 91, 6815–6818. doi: 10.1073/pnas.91.15.6815

IBGSC (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491, 711–716. doi: 10.1038/nature11543

Khademian, H., Giraut, L., Drouaud, J., and Mezard, C. (2013). Characterization of meiotic non-crossover molecules from *Arabidopsis thaliana* pollen. *Methods Mol. Biol.* 990, 177–190. doi: 10.1007/978-1-62703-333-6_18

Künzel, G. (1982). Differences between genetic and physical centromere distances in the case of two genes for male sterility in barley. *Theor. Appl. Genet.* 64, 25–29. doi: 10.1007/BF00303645

Künzel, G., Korzun, L., and Meister, A. (2000). Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. *Genetics* 154, 397–412.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*. arXiv:1303.3997v2.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, X., Li, L., and Yan, J. (2015). Dissecting meiotic recombination based on tetrad analysis by single-microspore sequencing in maize. *Nat. Commun.* 6:6648. doi: 10.1038/ncomms7648

Linde-Laursen, I. (1982). Linkage map of the long arm of barley chromosome-3 using C-bands and marker genes. *Heredity (Edinb).* 49, 27–35. doi: 10.1038/hdy.1982.62

Lu, P. L., Han, X. W., Qi, J., Yang, J. G., Wijeratne, A. J., Li, T., et al. (2012). Analysis of *Arabidopsis* genome-wide variations before and after meiosis and meiotic recombination by resequencing Landsberg erecta and all four products of a single meiosis. *Genome Res.* 22, 508–518. doi: 10.1101/gr.127522.111

Martin, A. C., Rey, M. D., Shaw, P., and Moore, G. (2017). Dual effect of the wheat Ph1 locus on chromosome synapsis and crossover. *Chromosoma* doi: 10.1007/s00412-017-0630-0. [Epub ahead of print].

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., et al. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature* 544, 427–433. doi: 10.1038/nature22043

Mascher, M., Muehlbauer, G. J., Rokhsar, D. S., Chapman, J., Schmutz, J., Barry, K., et al. (2013a). Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). *Plant J.* 76, 718–727. doi: 10.1111/tpj.12319

Mascher, M., Richmond, T. A., Gerhardt, D. J., Himmelbach, A., Clissold, L., Sampath, D., et al. (2013b). Barley whole exome capture: a tool for genomic research in the genus *Hordeum* and beyond. *Plant J.* 76, 494–505. doi: 10.1111/tpj.12294

Melamed-Bessudo, C., and Levy, A. A. (2012). Deficiency in DNA methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 109, E981–E988. doi: 10.1073/pnas.1120742109

Mirouze, M., Lieberman-Lazarovich, M., Aversano, R., Bucher, E., Nicolet, J., Reinders, J., et al. (2012). Loss of DNA methylation affects the recombination landscape in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5880–5885. doi: 10.1073/pnas.1120841109

Nilsson, N. O., and Pelger, S. (1991). The relationship between natural variation in chiasma frequencies and recombination frequencies in barley. *Hereditas* 115, 121–126. doi: 10.1111/j.1601-5223.1991.tb03545.x

Pedersen, S. (1988). Pollen competition in barley. *Hereditas* 109, 75–81. doi: 10.1111/j.1601-5223.1988.tb00185.x

Phillips, D., Jenkins, G., Macaulay, M., Nibau, C., Wnetrzak, J., Fallding, D., et al. (2015). The effect of temperature on the male and female recombination landscape of barley. *New Phytol.* 208, 421–429. doi: 10.1111/nph.13548

Phillips, D., Wnetrzak, J., Nibau, C., Barakate, A., Ramsay, L., Wright, F., et al. (2013). Quantitative high resolution mapping of HvMLH3 foci in barley pachytene nuclei reveals a strong distal bias and weak interference. *J. Exp. Bot.* 64, 2139–2154. doi: 10.1093/jxb/ert079

Ritz, K. R., Noor, M. A. F., and Singh, N. D. (2017). Variation in recombination rate: adaptive or not? *Trends Genet.* 33, 364–374. doi: 10.1016/j.tig.2017.03.003

Sall, T. (1990). Genetic control of recombination in barley. 2. Variation in linkage between marker genes. *Hereditas* 112, 171–178. doi: 10.1111/j.1601-5223.1990.tb00054.x

Sall, T., Flink, J., and Bengtsson, B. O. (1990). Genetic control of recombination in barley. 1. Variation in recombination frequency measured with inversion heterozygotes. *Hereditas* 112, 157–170. doi: 10.1111/j.1601-5223.1990.tb00053.x

Salome, P. A., Bomblies, K., Fitz, J., Laitinen, R. A., Warthmann, N., Yant, L., et al. (2012). The recombination landscape in *Arabidopsis thaliana* F2 populations. *Heredity (Edinb).* 108, 447–455. doi: 10.1038/hdy.2011.95

Sarigorla, M., Pe, M. E., Mulcahy, D. L., and Ottaviano, E. (1992). Genetic dissection of pollen competitive ability in maize. *Heredity (Edinb).* 69, 423–430. doi: 10.1038/hdy.1992.146

Sidhu, G. K., Fang, C., Olson, M. A., Falque, M., Martin, O. C., and Pawlowski, W. P. (2015). Recombination patterns in maize reveal limits to crossover homeostasis. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15982–15987. doi: 10.1073/pnas.1514265112

Smith, J. M., and Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genet. Res.* 23, 23–35. doi: 10.1017/S0016672300014634

Sybenga, J. (1966). Quantitative analysis of chromosome pairing and chiasma formation based on relative frequencies of M I configurations. 4. Interchange heterozygotes. *Genetica* 37, 199–206. doi: 10.1007/BF01547131

Taylor, D. R., and Ingvarsson, P. K. (2003). Common features of segregation distortion in plants and animals. *Genetica* 117, 27–35. doi: 10.1023/A:1022308414864

Walsh, N. E., and Charlesworth, D. (1992). Evolutionary interpretations of differences in pollen tube growth rates. *Q. Rev. Biol.* 67, 19–37. doi: 10.1086/417446

Wijnker, E., Velikkakam James, G., Ding, J., Becker, F., Klasen, J. R., Rawat, V., et al. (2013). The genomic landscape of meiotic crossovers and gene conversions in *Arabidopsis thaliana*. *eLife* 2:e01426. doi: 10.7554/eLife.01426

Wright, S. I., Foxe, J. P., DeRose-Wilson, L., Kawabe, A., Looseley, M., Gaut, B. S., et al. (2006). Testing for effects of recombination rate on nucleotide diversity in natural populations of *Arabidopsis* lyrata. *Genetics* 174, 1421–1430. doi: 10.1534/genetics.106.062588

Xu, Y., Zhu, L., Xiao, J., Huang, N., and McCouch, S. R. (1997). Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* 253, 535–545. doi: 10.1007/s0043800 50355

Yelina, N. E., Choi, K., Chelysheva, L., Macaulay, M., de Snoo, B., Wijnker, E., et al. (2012). Epigenetic remodeling of meiotic crossover frequency in *Arabidopsis thaliana* DNA methyltransferase mutants. *PLoS Genet.* 8:e1002844. doi: 10.1371/journal.pgen.1002844

Yelina, N. E., Ziolkowski, P. A., Miller, N., Zhao, X., Kelly, K. A., Munoz, D. F., et al. (2013). High-throughput analysis of meiotic crossover

frequency and interference via flow cytometry of fluorescent pollen in *Arabidopsis thaliana*. *Nat. Protoc.* 8, 2119–2134. doi: 10.1038/nprot. 2013.131

Ziolkowski, P. A., Berchowitz, L. E., Lambing, C., Yelina, N. E., Zhao, X., Kelly, K. A., et al. (2015). Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover remodelling via interference during *Arabidopsis* meiosis. *eLife* 4:e03708. doi: 10.7554/eLife.03708

Ziolkowski, P. A., Underwood, C. J., Lambing, C., Martinez-Garcia, M., Lawrence, E. J., Ziolkowska, L., et al. (2017). Natural variation and dosage of the HEI10 meiotic E3 ligase control *Arabidopsis* crossover recombination. *Genes Dev.* 31, 306–317. doi: 10.1101/gad.295501.116

3.3. Live cell CRISPR-imaging in plants reveals dynamic telomere movements

Dreissig, S. et al., 2017. Live cell CRISPR-imaging in plants reveals dynamic telomere movements. The Plant Journal, doi:10.1111/tpj.13601.

TECHNICAL ADVANCE

Live-cell CRISPR imaging in plants reveals dynamic telomere movements

Steven Dreissig[1] (iD), Simon Schiml[2], Patrick Schindele[2], Oda Weiss[1], Twan Rutten[1], Veit Schubert[1] (iD), Evgeny Gladilin[1], Michael F. Mette[1,†], Holger Puchta[2,*] (iD) and Andreas Houben[1,*] (iD)

[1]*Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, 06466 Seeland, Germany, and*
[2]*Botanical Institute, Karlsruhe Institute of Technology, POB 6980, 76049 Karlsruhe, Germany*

SUMMARY

Elucidating the spatiotemporal organization of the genome inside the nucleus is imperative to our understanding of the regulation of genes and non-coding sequences during development and environmental changes. Emerging techniques of chromatin imaging promise to bridge the long-standing gap between sequencing studies, which reveal genomic information, and imaging studies that provide spatial and temporal information of defined genomic regions. Here, we demonstrate such an imaging technique based on two orthologues of the bacterial clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR associated protein 9 (Cas9). By fusing eGFP/mRuby2 to catalytically inactive versions of *Streptococcus pyogenes* and *Staphylococcus aureus* Cas9, we show robust visualization of telomere repeats in live leaf cells of *Nicotiana benthamiana*. By tracking the dynamics of telomeres visualized by CRISPR–dCas9, we reveal dynamic telomere movements of up to 2 μm over 30 min during interphase. Furthermore, we show that CRISPR–dCas9 can be combined with fluorescence-labelled proteins to visualize DNA–protein interactions *in vivo*. By simultaneously using two dCas9 orthologues, we pave the way for the imaging of multiple genomic loci in live plants cells. CRISPR imaging bears the potential to significantly improve our understanding of the dynamics of chromosomes in live plant cells.

Keywords: CRISPR–dCas9, live cell imaging, telomeres, chromatin dynamics, nucleus, *Nicotiana benthamiana*, technical advance.

INTRODUCTION

The spatial and temporal organization of genomes is important for maintaining and regulating cell functions such as gene expression, DNA replication and repair, and the proper segregation of genetic material during cell division. Elucidating how the genome is spatiotemporally organised inside the nucleus is imperative to our understanding of how genes and non-coding sequences are regulated during development. Mapping the functional organization of the genome can be achieved by visualizing interactions between different genomic elements in living cells. Although fluorescence-tagged nuclear proteins can be readily imaged in living plant cells, the *in vivo* visualization of defined DNA sequences is technically difficult. Fluorescent *in situ* hybridization (FISH), a well-established tool to map DNA sequences, relies on fixed tissue samples and

cannot be used to visualize dynamic processes. Furthermore, FISH requires cell fixation and a DNA denaturation step that may result in an altered chromatin structure (Kozubek *et al.*, 2000; Boettiger *et al.*, 2016).

Live cell labelling of specific genomic loci has been achieved by the application of a directly repeated *lac* operator sequence and its detection with a GFP-lacI repressor protein (Kato and Lam, 2001). However, this system is based on the random insertion of an artificial sequence into the genome. Live imaging of endogenous genomic regions became possible with the application of fluorescence-tagged zinc-finger proteins. A zinc-finger GFP protein was designed to recognize a 9-bp sequence within the centromeric 180-bp tandem repeat of *Arabidopsis thaliana* (Lindhout *et al.*, 2007). Despite the numerous

uses of engineered zinc-finger proteins for genome editing, the potential of this technology has not yet been fully exploited for chromatin imaging. The discovery of the *Xanthomonas*-based DNA binding domain (Boch *et al.*, 2009), which can be engineered to bind specific DNA sequences, initiated the development of transcription activator-like effectors (TALEs) fused with fluorescence proteins (Ma *et al.*, 2013). Although fluorescently labelled TALEs were successfully used to trace genomic loci in non-plant organisms (reviewed in Chen *et al.*, 2016a), their application in plants has only recently been shown by Fujimoto *et al.*, (2016).

The discovery of the type-II clustered regularly interspaced short palindromic repeats (CRISPR) system derived from *Streptococcus pyogenes*, has revolutionized the field of targeted genome editing in eukaryotes (Jinek *et al.*, 2012). Cas9 nuclease-based genome engineering has become a routine technology for many plant species (reviewed in Pacher and Puchta, 2017); however, the full potential of this technology reaches far beyond the controlled induction of mutations. The Cas9 nuclease can be transformed by two point mutations into a site-specific DNA-binding protein, which can be fused with different protein domains. Thus, it should in principle be possible to target any kind of enzymatic activity to any genomic site of interest (Puchta, 2016a). Recently, nuclease-deficient derivatives (dCas9) were used to modify gene expression in many model organisms, including plants (Qi *et al.*, 2013; Piatek *et al.*, 2015). Furthermore, by fusing dCas9 with GFP, the CRISPR–dCas9 system has been used to label genomic loci in live mammalian cells (Chen *et al.*, 2013; Anton *et al.*, 2014). Multicolour CRISPR–dCas9 imaging became possible with the application of dCas9 orthologues from different bacterial species, like *Neisseria meningitidis* (NmCas9), *Streptococcus thermophilus* (St1Cas9), and *Staphylococcus aureus* (SaCas9) (Ma *et al.*, 2015; Chen *et al.*, 2016b). The discovery of the Cas9-like activities of the Cpf1 proteins derived from *Acidaminococcus* and *Lachnospiraceae* (Zetsche *et al.*, 2015) may further expand the palette of multicolour CRISPR–dCas9 imaging. More importantly, orthologues of *S. pyogenes* (Sp-Cas9), such as St1-Cas9 and Sa-Cas9,

have been confirmed to be functional in plants (Steinert *et al.*, 2015).

In the current study, we describe the development of CRISPR–dCas9 for live cell imaging in plants based on two Cas9 orthologues derived from *S. pyogenes* (Sp-dCas9) and *S. aureus* (Sa-dCas9). We demonstrate reliable imaging of telomere repeats in living cells of *Nicotiana benthamiana* and pave the way for the potential visualization of multiple genomic loci. Furthermore, we show that CRISPR–dCas9 can be combined with fluorescence-labelled proteins to investigate DNA–protein interactions *in vivo*.

RESULTS & DISCUSSION

CRISPR–dCas9 enables the visualization of tandem repeats in live plant cells

To establish live cell imaging by CRISPR–dCas9 in plants, we introduced a point mutation in the RuvC1 and HNH nuclease domains (D10A and H841A) in two Cas9 orthologues derived from *S. pyogenes* (Sp-Cas9) and *S. aureus* (Sa-Cas9), which were previously used for targeted mutagenesis in *A. thaliana* (Fauser *et al.*, 2014; Steinert *et al.*, 2015), rendering the Cas9 protein catalytically inactive. Multiple copies of fluorescence proteins, either eGFP or mRuby2, were fused to the C-terminal end of each dCas9 variant (Figure 1a).

To test the functionality of CRISPR–dCas9 for live cell imaging in plants, we imaged the telomeres of *Nicotiana benthamiana* in leaf cells. *Nicotiana* telomeres are composed of arrays of TTTAGGG repeats that are 60–160 kb in length (Fajkus *et al.*, 1995). These tandemly repeated DNA sequences allow the binding of many dCas9 proteins at the same locus by a single single-guide (sg) RNA sequence. To target telomeric repeats we constructed a 20-nucleotide sgRNA that was complementary to the TTTAGGG telomere sequence, starting with a 'G' at the 5′ end (sgRNA-telomere; Figure 1b). The 5′-G nucleotide was selected as the *A. thaliana* U6-26 promoter employed requires it to initiate transcription (Belhaj *et al.*, 2013).

We used infiltration by *Agrobacterium tumefaciens* to transiently express Sp-dCas9 and sgRNA-telomere in leaf cells. As a negative control, the same dCas9 construct was infiltrated without a specific sgRNA. After 2–4 days, bright
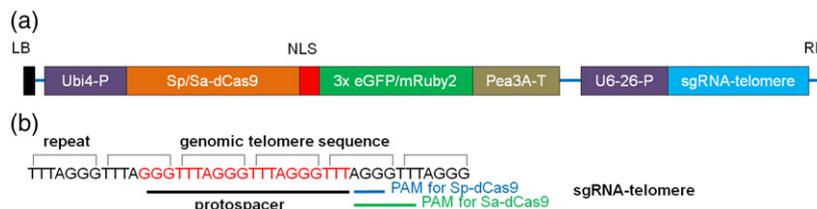
Figure 1. Structure of the CRISPR–dCas9 construct. (a) Transcription of Sp/Sa-dCas9 was initiated by the parsley ubiquitin 4 promoter and terminated by the pea 3A terminator. An SV40 NLS DNA sequence was used for nuclear localization of dCas9. Transcription of the sgRNA scaffold was initiated by the *Arabidopsis* ubiquitin 6 promoter. (b) Protospacer design for Sp-dCas9 and Sa-dCas9 to target telomere DNA sequence. Target sequence is shown in red. The NGG protospacer adjacent motif (PAM) for Sp-dCas9 is indicated in blue, whereas *NNGRRT* PAM for Sa-dCas9 is indicated in green.

fluorescence puncta were observed in addition to a weak nonspecific background labelling of the nucleus and in particular of the nucleolus (Figure 2a). A similar nonspecific labelling of the nucleolus caused by CRISPR–dCas9 was observed in previous studies (Chen *et al.*, 2013; Zhou *et al.*, 2017). In negative controls, only a weak nonspecific labelling of the nucleus was observed (Figure 2b). In live interphase nuclei, we detected an average of 21.75 telomere signals by CRISPR imaging (*n* = 50 nuclei). To confirm the telomere specificity of the fluorescence signals and to quantify the efficiency of dCas9 telomere labelling, we analysed the co-localization of dCas9 and telomeres by immunofluorescence and fluorescent *in situ* hybridization (FISH; Figure 2c–e). On average, 27 signals were detected by immunofluorescence labelling of dCas9, which amounts to 78% of all telomeres detected by FISH (Figure 2e, f). We observed an average of 35 telomere FISH signals, which indicates a certain degree of telomere clustering as the expected number of telomeres based on a

chromosome complement of 2*n* = 38 would be 76 in 2C nuclei (Appendix S1). Notably, a similar localization pattern was observed in wild-type leaf interphase nuclei (Appendix S2), although we detected a higher number of telomere signals (average = 53, *n* = 30). We then used dCas9 without introducing a specific sgRNA as a control, and detected an average of 42 signals (*n* = 30). As the detectable number of telomeres appears to be highly variable in *N. benthamiana* leaf nuclei, we conclude that mainly nuclei with clustered telomeres were visualized in our experimental system. A higher number of CRISPR–dCas9 signals were observed in fixed cells after immunofluorescence labelling compared to live cells, which might be a result of the improved detection efficiency of the GFP antibody. The intensity of individual hybridization signals most likely varied because of chromosome-specific differences in telomere repeat number and fusion of chromosome ends. Importantly, we observed a positive correlation ($\rho = 0.84$, $r^2 = 0.7$, $n = 30$)

Figure 2. Live imaging of telomeres by CRISPR–dCas9. (a) Sp-dCas9-mRuby and sgRNA -telomere were used for live imaging of telomeres in *N. benthamiana* leaf cells during interphase (*n* = 50). (b) As a negative control, the telomere sgRNA was omitted. (c–e) Immunofluorescence staining against Sp-dCas9-eGFP (c), combined with FISH against telomeres (d), and overlain to confirm co-localization. Nucleus is counterstained with DAPI (in blue) (e). (f) Whisker box plot showing the efficiency of Sp-dCas9 for telomere labelling (*n* = 50 nuclei). CRISPR live refers to the number of signals in live leaf nuclei, whereas CRISPR fixed refers to the number of signals in isolated nuclei after fixation. – sgRNA indicates the number of telomeres counted after transformation of dCas9 without the sgRNA-telomere. (g) Intensity plot showing a positive correlation between FISH (red) and CRISPR imaging (green) regarding the size and intensity of hybridization signals (indicated by the dotted line in panel e) (*n* = 30). Scale bars: 10 μm.
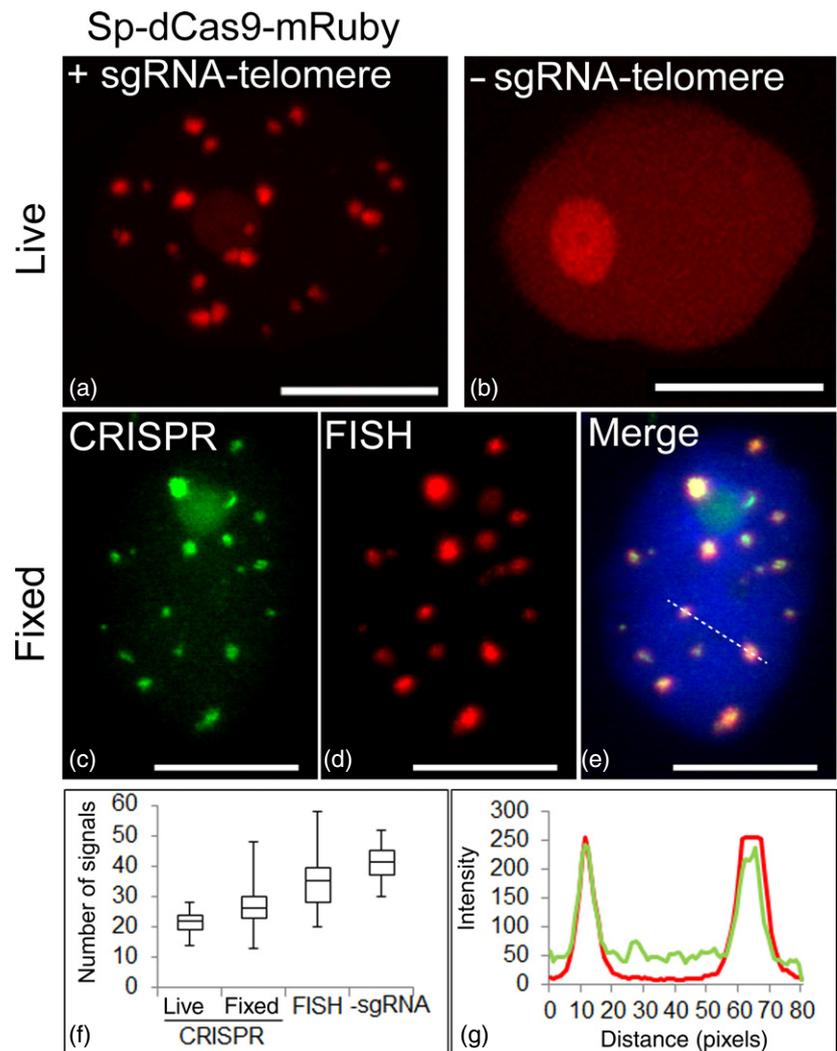
between FISH and CRISPR imaging regarding the intensity and size of hybridization signals (Figure 2g).

Imaging of telomeres by CRISPR–dCas9 reveals long-range telomere dynamics

We then attempted to explore the possibility of visualizing telomere movement in a live nucleus by CRISPR imaging. To visualize the nuclear envelope in addition to telomeres, we used the nuclear egress protein of the human cytome-galovirus pUL50 fused to GFP, which was previously shown to localize to the *N. benthamiana* nuclear envelope (Lamm *et al.*, 2016, Appendix S3). Individual nuclei were observed *in vivo* for a total of 30 min, and *z*-stacks were acquired in 1-min intervals. We observed stable fluorescence over the entire period of time.

To investigate whether there is a stable association of dCas9 with its target sequence or rapid turnover, we conducted fluorescence recovery after photobleaching (FRAP) experiments. After bleaching individual dCas9 clusters of three nuclei, we found no significant fluorescence recovery over a period of 30 min (Figure 3). This indicates a stable association of dCas9 with its target sequence, which is in agreement with previous reports showing an average target residence time of more than 3 h in mammalian cells (Ma *et al.*, 2016; Qin *et al.*, 2017).

Several dynamic subcellular movements were observed. First, entire nuclei showed movements in all three dimensions over time. Within those nuclei, however, telomeres tended to follow these movements, but also changed their position relative to each other (Appendix S4) and were localized in proximity to the nuclear envelope, which was shown by a mean normalized radial distance of telomeres of 0.74 (Figure 4a; Appendix S5). To quantify these dynamics, we tracked the spatial movements of telomeres over the entire period of time and measured their mean square displacement (MSD; Figure 4b; Appendices S6 and S7). By tracking single telomere clusters, we observed confined diffusion of telomere clusters as well as long-range movements, which results in a high standard deviation of the MSD. To reveal these variations in a representative nucleus, we calculated absolute changes in intertelomere distance over time (Figure 4c, d). Within 30 min, individual telomeres changed their distance from each other by up to ±2 μm (average diameter of the nucleus = 15.12 μm, *n* = 12). Compared with an average intertelomere distance of 8.1 μm, these changes can amount to a maximum of 24.7%. Similar observations were made previously in *A. thaliana* by labelling telomeres via fluorescent TALEs (Fujimoto *et al.*, 2016). In contrast to Fujimoto *et al.* (2016), however, we did not observe an active formation of telomere clusters. This might be related to differences in interphase telomere dynamics between *A. thaliana* root cells and *N. benthamiana* leaf cells rather than CRISPR–dCas9 having a negative effect on telomere
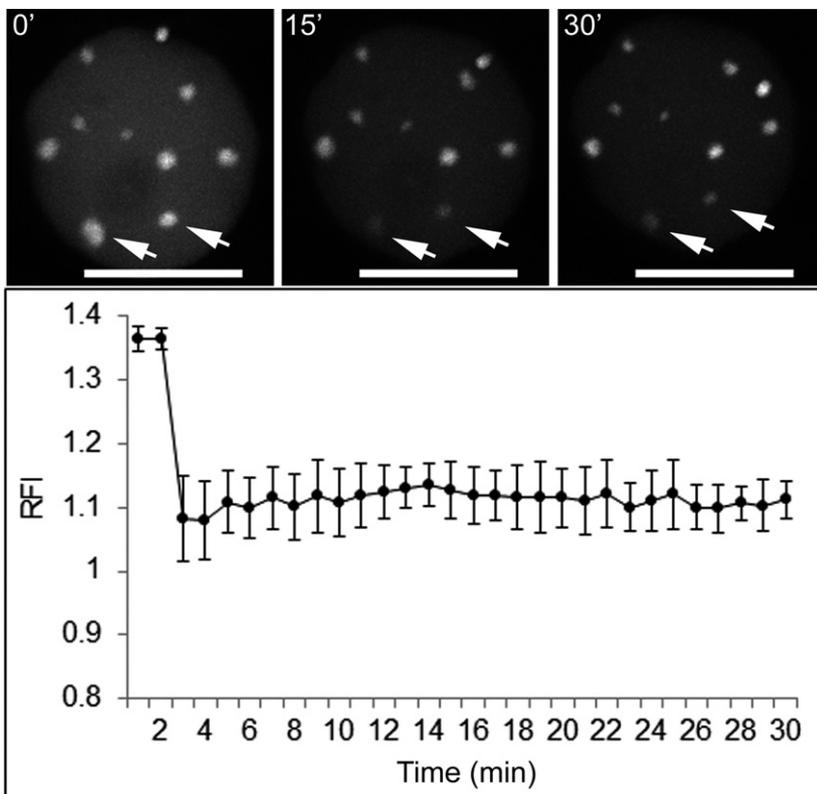
Figure 3. Fluorescence recovery after photobleaching (FRAP) analysis demonstrates a stable association of dCas9 with the target sequence during interphase. FRAP experiments were conducted on three individual nuclei. A region of interest was bleached (indicated by arrows) and the fluorescence intensity was compared with the background fluorescence intensity to determine the relative fluorescence intensity (RFI). Error bars represent standard deviations, based on three biological replicates. Scale bars: 10 μm.

Figure 4. CRISPR–dCas9 enables the 3D tracking of telomeres and reveals long-range movements in interphase nuclei. (a) Normalized radial distance (NDR) of telomeres of a representative nucleus. An NDR of 0 indicates localization in the centre of the nucleus, whereas an NDR of 1 indicates localization at the nuclear envelope. Error bars represent standard deviations, based on measurements conducted at different time points (1–30 min). Telomere number represents individual telomere signals in a live nucleus. (b) Mean square displacement (MSD) in μm^2 was measured in 12 live nuclei with a total of 181 telomere signals over a period of 30 min. Error bars represent standard deviations. (c) Heat map showing changes in intertelomere distance over a period of 30 min in a representative nucleus. Colours represent increased (green) and decreased (red) intertelomere distances, by up to 2 μm. (d) Simultaneous visualization of 3D telomere locations in the same nucleus as in (c) from all time points after rigid registration to a reference system of coordinates given by the first time point.
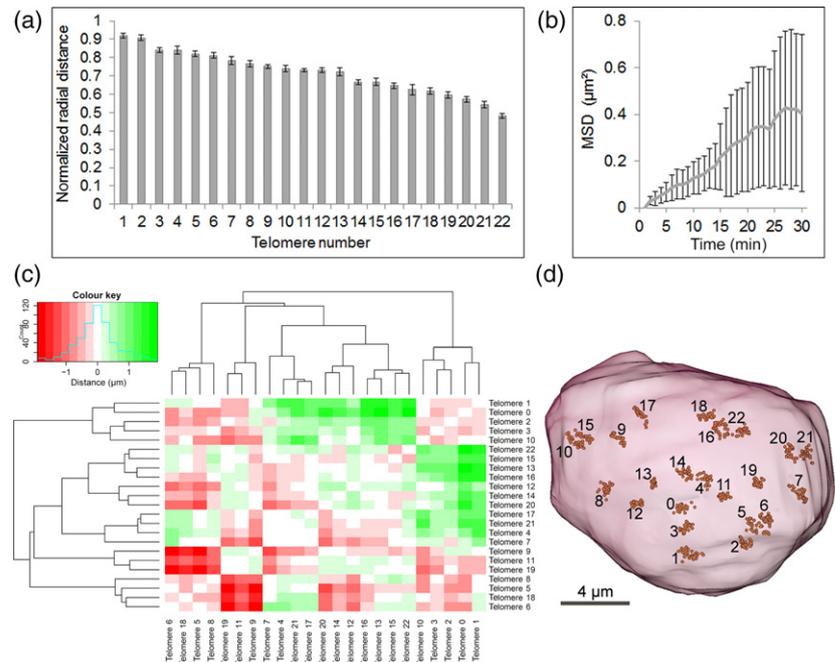
dynamics, as both MSD curves are similar. Telomere dynamics during interphase might be related to the transcription of telomeric tandem repeats (Koo *et al.*, 2016), telomerase activity (Schrumpfova *et al.*, 2016), or positional silencing by telomeres (Gottschling *et al.*, 1990; Nimmo *et al.*, 1994; Cryderman *et al.*, 1999). Similar long-range chromatin dynamics of specific interstitial chromosomal regions were previously described based on fixed *A. thaliana* cells (Schubert *et al.*, 2014). Our results reveal that long-range chromatin movements can occur over a short period of time, which we suggest is highly relevant for chromatin conformation capture studies (reviewed in Bonev and Cavalli, 2016) aiming to look at such interactions. We conclude that CRISPR–dCas9 is a robust system to reveal the dynamics of telomeres in live plant cells.

Visualization of DNA–protein dynamics at plant telomeres

As CRISPR–dCas9 is a tool to visualize specific DNA sequences, it can be combined with other methods, e.g. fluorescence-labelled proteins, to study the dynamics of DNA–protein interactions. As an example, we attempted to visualize telomeric DNA by CRISPR–dCas9 and the telomeric repeat binding protein 1 (TRB1) in live leaf cells of *N. benthamiana*. TRB1 was previously found to be located at plant telomeres and to interact with telomerase reverse transcriptase (TERT), although not all telomeres are bound by TRB1 (Dvorackova *et al.*, 2010; Schrumpfova *et al.*, 2014). Interestingly, telomeres in yeast, ciliates, mammals and plants may form 3′ overhangs; however, in plants, blunt-ended telomeres and telomeres with 3′ overhangs

may appear in the same cell, and only telomeres exhibiting 3′ overhangs are bound by TRB1.

We simultaneously expressed Sp-dCas9-mRuby, sgRNA-telomere and TRB1-GFP in leaf cells to visualize the dynamic relationship between the telomeric repeats of all chromosomes, and the proportion that are bound by TRB1. On average, we detected 30 CRISPR–dCas9 signals resembling telomeres, 26.3 (87.6%) of which were simultaneously bound by TRB1 (Figure 5). This indicates that most telomeres in *N. benthamiana* form 3′ overhangs and only a small proportion of blunt-ended telomeres are present during the interphase. Our results demonstrate that CRISPR–dCas9 can be used to visualize specific DNA sequences in combination with fluorescently tagged proteins interacting with those DNA sequences. We hypothesize that this principle can be expanded to investigate spatiotemporal gene expression patterns, e.g. by visualizing DNA sequences and transcription factors, as well as other DNA–protein interactions, such as the loading of specific histone variants to certain genomic regions (e.g. CENH3 and centromeric DNA).

Comparing the efficiency of two dCas9 orthologues for the imaging of telomeres

In addition to Sp-dCas9 derived from *Streptococcus pyogenes*, we used Sa-dCas9 derived from *Staphylococcus aureus* to visualize telomeres, compare their efficiency and potentially pave the way for the simultaneous imaging of multiple genomic loci in plants. The protospacer-adjacent motif (PAM) required by Sa-dCas9 (NNGRRT) allowed us to use the same sgRNA as for Sp-dCas9. In contrast to
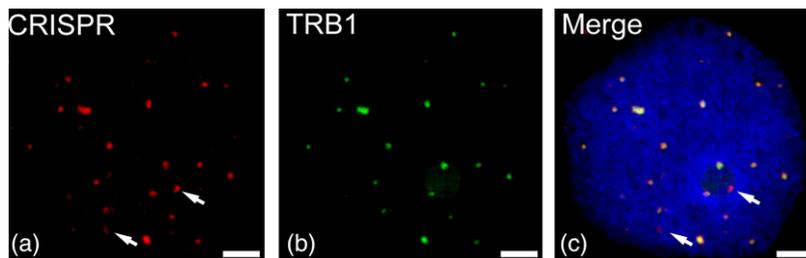
Figure 5. Simultaneous visualization of telomeric DNA by CRISPR–dCas9 and the GFP-tagged telomeric repeat binding protein 1 (TRB1). (a) Immunofluorescence staining against Sp-dCas9-mRuby2. (b) Immunofluorescence staining against TRB1-GFP. (c) Overlay showing almost complete co-localization, except for putative blunt-ended telomeres (indicated by arrows, nucleus is counterstained with DAPI (in blue). Scale bars: 2 μm.

Sp-dCas9, the sgRNA scaffold sequence of Sa-dCas9 as well as the size of the complex itself differs (Sa-dCas9 = 1064 amino acids versus Sp-dCas9 = 1380 amino acids), which could have an effect on its efficiency (Chen *et al.*, 2016b). In previous experiments, we showed that in plant cells Sp-Cas9 and Sa-Cas9 only form a complex with their respective sgRNAs, and not with the sgRNA of the other orthologue (Steinert *et al.*, 2015). Compared with Sp-dCas9 detecting 78% of all telomeres, using Sa-dCas9 we were able to detect 85.5% of all telomeres when both variants were analysed separately. We then used both variants Sp-dCas9-mRuby2 and Sa-dCas9-eGFP simultaneously to visualize telomeres and demonstrate the potential application of different dCas9 orthologues for multiple genomic loci. When combined, both variants showed almost complete co-localization, indicating no significant difference in their efficiency to detect telomeres (Figure 6). We conclude that both Sp-dCas9 and Sa-dCas9 can be used simultaneously to visualize tandem repeats in live plant cells. Thus, by using different Cas9 or Cpf1 orthologues fused to different fluorescence proteins, multidimensional live imaging in single cells might become a reality in the long run (Puchta, 2016b). For tandem repetitive sequences, a single sgRNA may be sufficient for CRISPR imaging, whereas labelling of non-repetitive loci (spanning a 5-kb region) may require the simultaneous expression of at least 30 sgRNAs (Chen *et al.*, 2013). More recently, an entire human chromosome was visualized with CRISPR imaging by using at least 485 non-repetitive sgRNAs at the same time (Zhou *et al.*, 2017). An alternative approach is to tether fluorescent RNA binding proteins to the sgRNA through aptamer fusions, thereby transforming the sgRNA into a scaffold RNA that contains information about the target locus and the type of fluorescence (Shao *et al.*, 2016). These recent developments may further improve CRISPR imaging in plants and potentially enable us to visualize single genomic loci.

EXPERIMENTAL PROCEDURES

T-DNA construction

All constructs are based on our previously described vector pCAS9-TPC (Fauser *et al.*, 2014). For the two Cas9 orthologues from *Streptococcus pyogenes* (Sp) and *Staphylococcus aureus* (Sa), respective dCas9 versions were generated by two consecutive rounds of site-directed mutagenesis, thereby introducing the

two point mutations (D10A/H840A for Sp-dCas9 and D10A/N580A for Sa-dCas9). Plasmids encoding for eGFP (pSiM24-eGFP) and mRuby2 (pcDNA3-mRuby2) were obtained from Addgene (http://www.addgene.com). dCas9 and fluorescence protein (FP) sequences were generated with primers containing homologous flanks for subsequent Gibson Assembly cloning into the pCAS9-TP backbone (for a full list of primers, see Appendix S8). The stop codons of the dCas9 and the first two FP sequences were removed to generate a continuous open reading frame (ORF), harbouring the respective dCas9 orthologue followed by a threefold fusion of the appropriate FP sequence. The dCas9 sequence and the FP fusion as well as the single FP sequences were linked via a GS-rich linker, respectively. Protospacers were allowed to self-anneal and the resulting 4-bp overhangs were used for subsequent ligation into the respective pChimera vector via *Bbs*I restriction sites. The customised RNA chimeras were then ligated into the respective dCas9 vectors via *Mlu*I restriction sites. The Cas9 constructs developed in this study are available on request to HP.

Protospacer design

Protospacer sequences were selected based on the respective PAM sequence of each dCas9 orthologue, namely SpdCas9 and SadCas9, and synthesized as oligonucleotides with appropriate 4-bp 5′ overhangs for cloning into the respective pChimera vector. The telomere-specific protospacer (5′-GGGTTTAGGGTTTAGGGTTT-3′) is based on the Arabidopsis-type telomere repeat sequence 5′-(TTTAGGG)(n)-3′. As a result of the presence of both Sp-dCas9 (5′-NGG-3′) and Sa-dCas9 (5′-NNGRRT-3′) PAM sequences in the telomere repeat sequence, both variants were used to label telomeres, which allowed us to compare these two orthologues.

Transient transformation of *N. benthamiana*

All CRISPR–dCas9 constructs, TRB1-GFP (Schrumpfova *et al.*, 2014) and pUL50-GFP (Lamm *et al.*, 2016) were separately transformed into *Agrobacterium tumefaciens* strain GV3101 by electroporation. Agrobacteria containing Sp-dCas9, Sa-dCas9 and pUL50-GFP were cultured in YEB medium (5 g l^{-1} beef extract, 1 g l^{-1} yeast extract, 5 g l^{-1} peptone, 5 g l^{-1} sucrose, 300 mg l^{-1} MgSO4, 20 g l^{-1} agar) containing spectinomycin (100 μg ml^{-1}) and rifampicin (50 μg ml^{-1}). For TRB1-GFP, Agrobacteria were cultured in YEB medium containing kanamycin (100 μg ml^{-1}) and rifampicin (50 μg ml^{-1}). The transient transformation of *N. benthamiana* leaf cells was performed as described in Phan and Conrad (2016). For the transformation of multiple constructs, bacterial cultures with an OD_{600} between 1.0 and 1.3 were mixed in a 1:1 ratio prior to transformation. Plants were analysed 2–4 days after infiltration.

Immunofluorescence analysis and fluorescence *in situ* hybridization

Two or three days after leaf infiltration, nuclei were extracted by chopping a 1-cm² piece of leaf tissue in 1 ml of chromosome

Figure 6. Comparison of Sa-dCas9 and Sp-dCas9. Telomeres were visualized by the simultaneous application of two dCas9 orthologues (Sa-dCas9 and Sp-dCas9). (a) Immunofluorescence staining against Sa-dCas9-eGFP. (b) Immunofluorescence staining against Sp-dCas9-mRuby2. (c) Overlay showing complete co-localization. Nucleus is counterstained with DAPI (in blue). (d) Quantification of the number of telomere signals observed by two different dCas9 orthologues ($n = 18$). Scale bars: 10 μm.
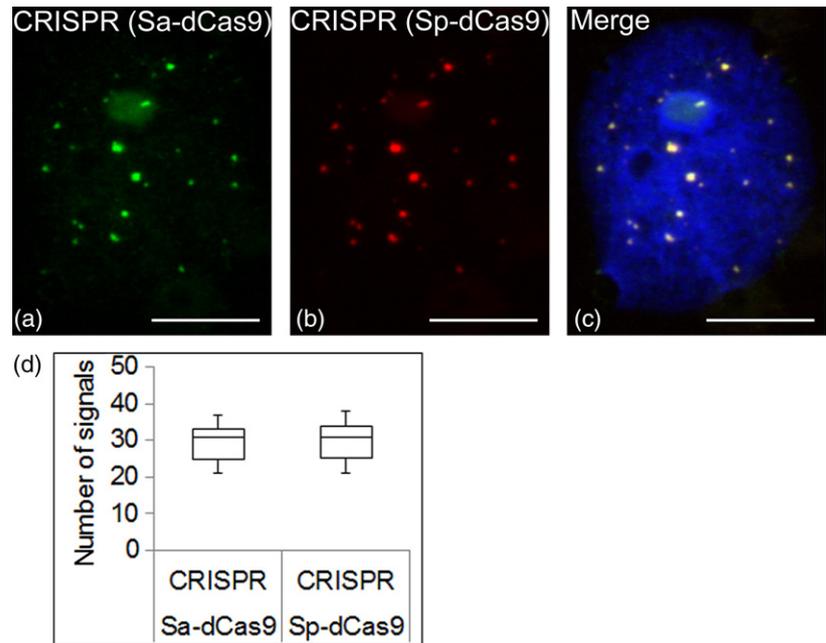
isolation buffer (Dolezel *et al.*, 2007) using a razor blade followed by filtration through a 35-μm nylon mesh and subsequent centrifugation onto a microscopic slide at 400 rpm for 5 min (Shandon CytoSpin3, https://gmi-inc.com). To confirm the specificity of each dCas9 construct, we conducted immunofluorescence staining against eGFP and mRuby2 in combination with fluorescence *in situ* hybridization (immuno-FISH) against telomeres. Immuno-FISH was performed as described by Ishii *et al.* (2015). eGFP was detected with a polyclonal GFP antibody (GFP antibody Dylight 488; Rockland, https://www.rockland-inc.com) in a 1:2500 dilution. mRuby2 was detected with a primary RFP antibody [RFP antibody (5F8), 1:1000 dilution; Chromotek, http://www.chromotek.com] generated in rats followed by an anti-rat secondary antibody (ab96889; abcam, http://www.abcam.com). To detect telomeres via FISH, we used 5′-Cy5 labelled oligonucleotides composed of the same DNA sequence as the respective protospacer (sgRNA-telomere, 5′-GGGTTTAGGGTTTAGGGTTT-3′). A final probe concentration of 0.33 μM was used. The correct telomeric localization of our FISH probe was validated by testing on *N. benthamiana* chromosomes (Appendix S1) prepared from flower buds using a protocol described by Sanchez Moran *et al.* (2001).

Microscopic analyses

To analyse co-localization between CRISPR–dCas9 imaging and FISH signals, images were acquired with an epifluorescence microscope (BX61; Olympus, https://www.olympus.com) using a cooled charge coupled device (CCD) camera (Orca ER; Hamamatsu, http://www.hamamatsu.com), and analysed with IMAGEJ. A total of 50 nuclei were analysed by immuno-FISH to determine the efficiency of dCas9 to detect telomeres. The number of *in vivo* dCas9 signals was counted in 50 live nuclei to determine the *in vivo* labelling efficiency. To analyse the co-localization of telomeres visualized by Sp-dCas9 and TRB1, a total of 43 nuclei were analysed by epifluorescence microscopy. Structured illumination microscopy (SIM) was applied to a representative sample using a 63×/1.4Oil Plan-Apochromat objective of an Elyra PS.1 microscope system with ZEN software (Carl Zeiss, https://www.zeiss.com).

Image stacks were captured separately for each fluorochrome using appropriate excitation and emission filters. Maximum intensity projections were generated from the stacks of SIM sections through the specimens in ZEN (3D rendering based on SIM image stacks was carried out using IMARIS 8.0; Bitplane, http://www.bitplane.com).

For live cell imaging of telomeres, fluorescence signals were analysed 2–4 days after infiltration with *A. tumefaciens* (see transient transformation of *N. benthamiana*) by a LSM780 (Carl Zeiss). Infiltrated leaf areas were cut and mounted onto a microscopic slide. The distribution of fluorescence signals within the nucleus was recorded as z-stacks ($n = 50$ nuclei). For a co-distribution analysis, probes were excited with dual 488- and 561-nm laser lines in combination with a 488/561-nm beam splitter. eGFP emission was detected over a range of 490–540 nm, and mRuby2 emission was detected over a range of 570–620 nm. Photospectrometric analysis of the fluorescence signal by means of the META detector confirmed the identity of GFP and mRuby. The turnover of Sp-dCas9-eGFP telomeric signals was investigated by FRAP analysis. After two pre-scans a region of interest of variable size was bleached. To achieve appropriate bleaching, the 488-nm laser line was set at 100% power with 25 iterations at scan speed 7. Fluorescence intensity was followed over 30 min in 1-min intervals.

Tracking of telomere signals and 3D image analysis

Telomere tracking based on time-lapse z-stacks was conducted with IMARIS 8.0 (Bitplane). Brightness was manually adjusted to detect all telomere clusters. Tracking was performed using the autoregression motion algorithm, with a maximum distance of 20 μm and a maximum gap size of 3. Afterwards, the coordinates (x, y, z) of each spot at all time points were used to quantify telomere movements. Intertelomere distances were calculated for all telomeres of a representative nucleus based on differences in distance between time point 1 and time point 30. For this purpose, an intertelomere distance matrix was generated for time point 1 (matrix1) and time point 30 (matrix30).

We then calculated the change in intertelomere distance by subtracting matrix1 from matrix30. The resulting matrix was then visualized as a heat map generated in RSTUDIO using the HEATMAP2 function of the GPLOTS package. Distances are presented in µm and visualized by two different colours, indicating an increase (green) or decrease (red) of intertelomere distance over time.

The 3D stacks of 12 live nuclei were semi-automatically segmented and triangulated surfaces of nuclear boundaries were generated in AMIRA 4.1 (Mercury Computer Systems, https://www.mrcy.com). To account for relative nuclear movements (i.e. translations, rotations), 3D point clouds of telomere mass centres from subsequent time points ($t > 0$) were rigidly registered to the reference system of coordinates given by the first time point ($t = 0$) using absolute orientation quaternions (Horn, 1987). To characterize the intranuclear telomere motion, the MSD of telomeres relative to their initial position ($t = 0$) was calculated as

$$\mathrm{MSD}(t) = \frac{1}{N}\sum_{i=1}^{N}\left(R_i(t) - R_i(0)\right)^2 \qquad (1)$$

where $R_i(t)$ denotes the radius vector of the i-th registered telomere in the reference system of coordinates at time point $t > 0$. The intranuclear position of telomeres was quantified in three representative nuclei by the normalized radial distance (NRD):

$$\mathrm{NRD} = TN/BN \qquad (2)$$

where TN and BN are the Euclidean distances between the nuclear envelope (N), the telomere (T) mass centres and the intersection point of the N–T line with the nuclear envelope surface (B), respectively. Accordingly, small NRD values indicate a nuclear-central telomere location, whereas values close to 1 correspond to the nuclear periphery.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors have no conflicts of interest to report.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Telomere FISH on *N. benthamiana* chromosomes.

Appendix S2. Telomere FISH on *N. benthamiana* wild-type interphase nucleus.

Appendix S3. Live interphase nucleus of *N. benthamiana* showing telomeres (Sp-dCas9-mRuby) and nuclear envelope (pUL50-GFP).

Appendix S4. Dynamic imaging of telomeres by CRISPR–dCas9.

Appendix S5. 3D telomere localization.

Appendix S6. Telomere tracking.

Appendix S7. Registration of telomere movements to reference system.

Appendix S8. Primers used for T-DNA construction.

REFERENCES

Anton, T., Bultmann, S., Leonhardt, H. and Markaki, Y. (2014) Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. *Nucleus*, **5**, 163–172.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S. and Nekrasov, V. (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods*, **9**, 39.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, **326**, 1509–1512.

Boettiger, A.N., Bintu, B., Moffitt, J.R., Wang, S., Beliveau, B.J., Fudenberg, G., Imakaev, M., Mirny, L.A., Wu, C.T. and Zhuang, X. (2016) Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature*, **529**, 418–422.

Bonev, B. and Cavalli, G. (2016) Organization and function of the 3D genome. *Nat. Rev. Genet.*, **17**, 661–678.

Chen, B., Gilbert, L.A., Cimini, B.A. *et al.* (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, **155**, 1479–1491.

Chen, B., Guan, J. and Huang, B. (2016a) Imaging specific genomic DNA in living cells. *Annu. Rev. Biophys.* **45**, 1–23.

Chen, B., Hu, J., Almeida, R., Liu, H., Balakrishnan, S., Covill-Cooke, C., Lim, W.A. and Huang, B. (2016b) Expanding the CRISPR imaging toolset with *Staphylococcus aureus* Cas9 for simultaneous imaging of multiple genomic loci. *Nucleic Acids Res.* **44**, e75.

Cryderman, D.E., Morris, E.J., Biessmann, H., Elgin, S.C.R. and Wallrath, L.L. (1999) Silencing at *Drosophila* telomeres: nuclear organization and chromatin structure play critical roles. *EMBO J.* **18**, 3724–3735.

Dolezel, J., Greilhuber, J. and Suda, J. (2007) Estimation of nuclear DNA content in plants using flow cytometry. *Nat. Protoc.*, **2**, 2233–2244.

Dvorackova, M., Rossignol, P., Shaw, P.J., Koroleva, O.A., Doonan, J.H. and Fajkus, J. (2010) AtTRB1, a telomeric DNA-binding protein from *Arabidopsis*, is concentrated in the nucleolus and shows highly dynamic association with chromatin. *Plant J.* **61**, 637–649.

Fajkus, J., Kovarik, A., Kralovics, R. and Bezdek, M. (1995) Organization of telomeric and subtelomeric chromatin in the higher-plant *Nicotiana tabacum*. *Mol. Gen. Genet.* **247**, 633–638.

Fauser, F., Schiml, S. and Puchta, H. (2014) Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J.* **79**, 348–359.

Fujimoto, S., Sugano, S.S., Kuwata, K., Osakabe, K. and Matsunaga, S. (2016) Visualization of specific repetitive genomic sequences with fluorescent TALEs in *Arabidopsis thaliana*. *J. Exp. Bot.* **67**, 6101–6110.

Gottschling, D.E., Aparicio, O.M., Billington, B.L. and Zakian, V.A. (1990) Position effect at *Saccharomyces cerevisiae* telomeres – reversible repression of Pol-II transcription. *Cell*, **63**, 751–762.

Horn, B.K.P. (1987) Closed-form solution of absolute orientation using unit quaternions. *J. Opt. Soc. Am. A*, **4**, 629–642.

Ishii, T., Sunamura, N., Matsumoto, A., Eltayeb, A.E. and Tsujimoto, H. (2015) Preferential recruitment of the maternal centromere-specific histone H3 (CENH3) in oat (*Avena sativa* L.) x pearl millet (*Pennisetum glaucum* L.) hybrid embryos. *Chromosome Res.* **23**, 709–718.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.

Kato, N. and Lam, E. (2001) Detection of chromosomes tagged with green fluorescent protein in live *Arabidopsis thaliana* plants. *Genome Biol.* **2**, research 0045.1–0045.10

Koo, D.H., Zhao, H.N. and Jiang, J.M. (2016) Chromatin-associated transcripts of tandemly repetitive DNA sequences revealed by RNA-FISH. *Chromosome Res.*, **24**, 467–480.

Kozubek, S., Lukasova, E., Amrichova, J., Kozubek, M., Liskova, A. and Slotova, J. (2000) Influence of cell fixation on chromatin topography. *Anal. Biochem.*, **282**, 29–38.

Lamm, C.E., Link, K., Wagner, S., Milbradt, J., Marschall, M. and Sonnewald, U. (2016) Human cytomegalovirus nuclear egress proteins ectopically expressed in the heterologous environment of plant cells are strictly targeted to the nuclear envelope. *Viruses*, **8**, 73.

Lindhout, B.I., Fransz, P., Tessadori, F., Meckel, T., Hooykaas, P.J. and van der Zaal, B.J. (2007) Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. *Nucleic Acids Res.*, **35**, e107.

Ma, H., Reyes-Gutierrez, P. and Pederson, T. (2013) Visualization of repetitive DNA sequences in human chromosomes with transcription activator-like effectors. *Proc. Natl. Acad. Sci. USA*, **110**, 21048–21053.

Ma, H., Naseri, A., Reyes-Gutierrez, P., Wolfe, S.A., Zhang, S. and Pederson, T. (2015) Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc. Natl. Acad. Sci. USA*, **112**, 3002–3007.

Ma, H., Tu, L.C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D. and Pederson, T. (2016) CRISPR–Cas9 nuclear dynamics and target recognition in living cells. *J. Cell Biol.*, **214**, 529–537.

Nimmo, E.R., Cranston, G. and Allshire, R.C. (1994) Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes. *EMBO J.*, **13**, 3801–3811.

Pacher, M. and Puchta, H. (2017) From classical mutagenesis to nuclease-based breeding – directing natural DNA repair for a natural end-product. *Plant J.* **90**, 819–833.

Phan, H.T. and Conrad, U. (2016) Plant-based vaccine antigen production. *Methods Mol. Biol.*, **1349**, 35–47.

Piatek, A., Ali, Z., Baazim, H., Li, L., Abulfaraj, A., Al-Shareef, S., Aouida, M. and Mahfouz, M.M. (2015) RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnol. J.*, **13**, 578–589.

Puchta, H. (2016a) Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. *Curr. Opin. Plant Biol.*, **36**, 1–8.

Puchta, H. (2016b) Using CRISPR/Cas in three dimensions: towards synthetic plant genomes, transcriptomes and epigenomes. *Plant J.* **87**, 5–15.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. and Lim, W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173–1183.

Qin, P., Parlak, M., Kuscu, C., Bandaria, J., Mir, M., Szlachta, K., Singh, R., Darzacq, X., Yildiz, A. and Adli, M. (2017) Live cell imaging of low- and non-repetitive chromosome loci using CRISPR–Cas9. *Nat. Commun.*, **8**, 14725.

Sanchez Moran, E., Armstrong, S.J., Santos, J.L., Franklin, F.C. and Jones, G.H. (2001) Chiasma formation in *Arabidopsis thaliana* accession Wassileskija and in two meiotic mutants. *Chromosome Res.* **9**, 121–128.

Schrumpfova, P.P., Vychodilova, I., Dvorackova, M., Majerska, J., Dokladal, L., Schorova, S. and Fajkus, J. (2014) Telomere repeat binding proteins are functional components of *Arabidopsis* telomeres and interact with telomerase. *Plant J.* **77**, 770–781.

Schrumpfova, P.P., Schorova, S. and Fajkus, J. (2016) Telomere- and telomerase-associated proteins and their functions in the plant cell. *Front. Plant Sci.* **7**, 851.

Schubert, V., Rudnik, R. and Schubert, I. (2014) Chromatin associations in *Arabidopsis* interphase nuclei. *Front. Genet.* **5**, 389.

Shao, S., Zhang, W., Hu, H., Xue, B., Qin, J., Sun, C., Sun, Y., Wei, W. and Sun, Y. (2016) Long-term dual-color tracking of genomic loci by modified sgRNAs of the CRISPR/Cas9 system. *Nucleic Acids Res.*, **44**, e86.

Steinert, J., Schiml, S., Fauser, F. and Puchta, H. (2015) Highly efficient heritable plant genome engineering using Cas9 orthologues from *Streptococcus thermophilus* and *Staphylococcus aureus*. *Plant J.* **84**, 1295–1305.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O. *et al.* (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR–Cas system. *Cell*, **163**, 759–771.

Zhou, Y., Wang, P., Tian, F., Gao, G., Huang, L., Wei, W. and Xie, X.S. (2017) Painting a specific chromosome with CRISPR/Cas9 for live-cell imaging. *Cell Res.*, **27**, 298–301.

4. Discussion

4.1. Why single cells?

Single-cell analyses differ from bulk-cell, whole organism, or population analyses in many technical aspects and in the scale at which new insights into biological processes can be made. Pioneering work in the field of single-cell analysis was conducted using microscopy-based methods that include a variety of protein or DNA staining approaches, and recently developed single-cell sequencing methods allowed interrogation of DNA sequence variation and RNA quantity in individual cells. These methodologies elucidated many fundamental biological aspects including cell division, protein localization, transcription, gene mapping, and DNA sequence variation (Cremer & Cremer, 2001; Dewitte & Murray, 2003; Sekar & Periasamy, 2003; Haraguchi, 2002; Aliyeva-Schnorr *et al*, 2015; Zong *et al*, 2012). The general aim was to gain mechanistic insight into such fundamental processes and to understand how individual cells contribute to form specific tissues or whole multicellular organisms. In plant research, progress in the field of genome sequencing has revealed the linear order of the genomes of many species, yet little is known about how different structures, e.g. interphase chromatin dynamics or meiotic chromosomes, contribute to certain functions such as transcription or meiotic recombination. These functions have significant implications for plant development and evolution. Single-cell analyses hold promise to help us gain new insights into these phenomena.

In the current study, the aim was to establish novel single-cell analysis methodologies based on single-cell genome sequencing and live cell imaging to allow to better understand the segregation and dynamics of specific genomic regions during meiosis and interphase. My work, although mostly technical, can be discussed in two ways. First, I will discuss its technical novelty and performance compared to other approaches. Secondly, I will discuss new insights that were made using these methodologies.

4.2. Analysis of meiotic recombination and segregation distortion in single pollen nuclei

Significant efforts are being directed towards the manipulation of meiotic recombination frequency and distribution in crops to unlock currently inaccessible genetic diversity in low-recombining regions of the genome. However, future attempts to change meiotic recombination frequency will require efficient methods to measure recombination. The aim of the current study was to develop single-cell analysis methodologies to measure meiotic recombination in individual pollen nuclei.

In order to develop a strategy for high-throughput analysis of meiotic recombination in pollen nuclei, I first tested different methods for isolation of pollen nuclei. Pollen grains are surrounded by a rigid cell wall which makes it difficult to isolate their DNA (Goss, 1968; Shi *et al*, 2015). These effects are exacerbated when handling single pollen because failure to isolate DNA would stop all downstream analyses. I therefore tested several methods, including chemical lysis (Chen *et al*, 2008; Gole *et al*, 2013), freeze-thaw cycles to disrupt the cell wall, osmotic shock to break the cell wall (Petersen *et al*, 1996), and mechanical disruption through small metallic beads (De Storme & Geelen, 2011). Although all of these methods can be used to isolate pollen DNA, there are several advantages and disadvantages associated with them. For example, chemical lysis requires strong alkaline conditions that subsequently need to be neutralized which increases the number of times contaminations can be introduced. Mechanical disruption of the pollen cell wall through metallic beads seemed to be the most suitable method because large amounts of nuclei could be isolated and individually sorted into reaction tubes via FACS (Dreissig *et al*, 2015).

Subsequently, two different methods were used for whole genome amplification, namely MDA and displacement-DOP-PCR (PicoPLEX). The major difference is that pollen DNA amplified via PicoPLEX was simultaneously tagged with adapter sequences for Illumina sequencing whereas pollen DNA amplified through MDA was analysed by allele-specific PCR. Both approaches were used to measure meiotic recombination but at different depth. With MDA-based amplification followed by allele-specific PCR, it was possible to analyse meiotic recombination frequency along barley chromosome 3H with 25 single-nucleotide-polymorphism (SNP) markers. Amplification errors were interrogated by

checking for heterozygous allele calls, which should be absent in haploid pollen nuclei, and by including a pure homozygous parental control which should not show any recombination events between the two parents unless introduced by amplification errors. Although no amplification errors were found in the parental control, 1.92% (24 out of 1250) of all allele-specific PCR reactions carried out with pollen nuclei appeared to contain heterozygous samples which was most likely caused by amplification errors or genotyping errors (Dreissig *et al*, 2015). However, this percentage seems to be comparable to previous studies where 0.6% and <1% heterozygous allele calls were reported (Wang *et al*, 2012; Li *et al*, 2015).

When displacement-DOP-PCR followed by whole-genome sequencing was used, heterozygous allele calls were removed for two reasons. First, a large number of SNPs was analyzed which allowed us to compare neighboring markers. Second, consensus genotypes were derived by aggregating information in 1 Mbp bins which effectively results in a lower resolution but higher confidence in allele calling. This resolution is higher than what can be achieved by conventional microscopy-based methods, such as the analysis of chiasmata during metaphase I by light microscopy (~15-20 Mbp, assuming a chromosome size of 8-10 μm like in barley and a optical resolution limit at 250 nm), and similar to high resolution microscopy analysis of pachytene chromosomes (~1 Mbp, assuming a synaptonemal complex length of 60-200 μm and a resolution limit at 100 nm) (Higgins *et al*, 2012; Phillips *et al*, 2012, 2013). However, one disadvantage of this pollen-sequencing approach is that small gene conversion events (<80 kbp, (Wijnker *et al*, 2013)) cannot be detected.

After errors introduced during amplification or genotyping were accounted for, both approaches were used to measure meiotic recombination in comparison to a DH population of the same genotype that was analysed through genotyping-by-sequencing (GBS, (International Barley Genome Sequencing Consortium *et al*, 2012). No differences between pollen and DH population were found regarding number and distribution of recombination events (Dreissig *et al*, 2015). This shows that both approaches are capable of producing reliable recombination measurements, with the potential to outperform cytological analyses and whole population analyses in terms of resolution and sample throughput.

Finally, applying these methodologies also enabled us to gain some new insights into meiotic recombination patterns and segregation. For example, the existence of class I and class II crossovers (interference sensitive vs insensitive) was not confirmed experimentally in barley yet, and our study, along with others (Phillips *et al*, 2013, 2015), provides supporting evidence for two crossover classes. Furthermore, pollen nuclei sequencing enabled us to analyse segregation distortion throughout the whole barley genome directly after meiosis and compare these data to segregation distortion in DH plants. Interestingly, we were able to provide experimental evidence that segregation distortion is almost absent in pollen whereas strong effects were observed in DH plants (Dreissig *et al*, 2017). This is in agreement with related studies (Bélanger *et al*, 2016b, 2016a; Manninen, 2000), but data derived from pollen was not available yet. It is therefore reasonable to conclude that segregation distortion is not a result of meiosis but rather caused by selection on linked loci during microspore culture, embryo development, plant regeneration, or spontaneous diploidization for DH production.
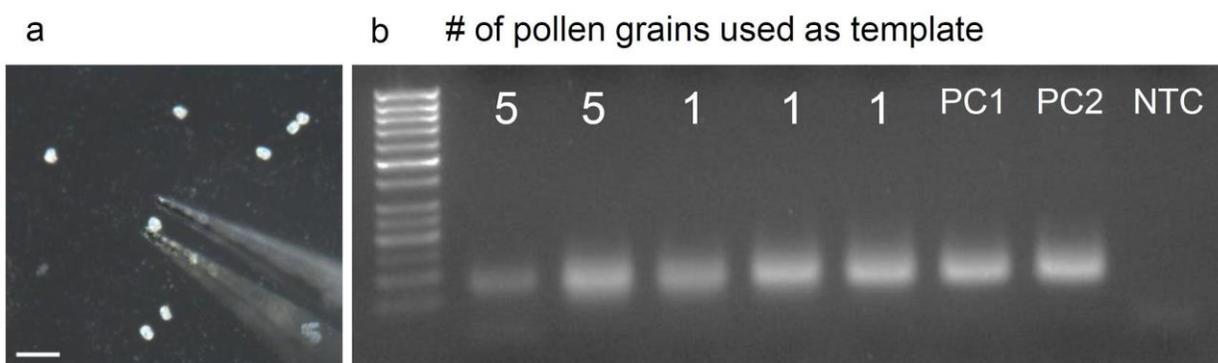
Figure 6 Manual pollen collection and PCR following chemical lysis of pollen.
(a) Pollen grains of barley were manually collected using a pair of forceps and a stereo microscope. **(b)** Chemical lysis was conducted as described in Chen (2008) followed by PCR using 5 or 1 pollen as template. A high-copy retrotransposon (*cereba*) of barley was amplified. Two positive controls (PC1 and PC2) with varying amounts of genomic DNA (40 ng vs 4 ng, respectively) and a no-template control (NTC) were included. The experiment shows that high-copy DNA sequences can readily be amplified from single pollen. However, manual collection of pollen grains is not preferred as the risk of introducing DNA contamination is high and because it is a time-consuming task.

4.3. Development of CRISPR-Cas9 for live cell imaging of defined genomic regions in plants

The field of live cell imaging of DNA sequences has emerged during the last two decades as an exciting new possibility to investigate the dynamics of genes and non-coding sequences during interphase and cell division. With the development of engineered DNA-binding proteins tagged with fluorescent proteins, such as the lactose operator/inhibitor sequence, ZFPs, and TALEs, significant progress was made (Kato & Lam, 2001; Lindhout *et al*, 2007; Fujimoto *et al*, 2016). More recently, CRISPR-Cas9 was discovered for genome editing and rapidly repurposed as a new live cell imaging system for DNA and RNA in mammalian cells by introducing two point mutations (D10A and H841A) which rendered the Cas9 nuclease deficient (dCas9) (Chen *et al*, 2013; Ma *et al*, 2015; Nelles *et al*, 2016).

The aim of the current study was to adapt CRISPR-Cas9 for live cell imaging in plants, provide experimental evidence for its performance, and demonstrate what new insights can be obtained using this method.

Due to previous efforts by our collaborators, who developed several CRISPR-Cas9 orthologues for genome editing in plants (Fauser *et al*, 2014; Schiml *et al*, 2014; Steinert *et al*, 2015), we were able to use Cas9 variants that were already codon optimized and functionally validated for dicotyledonous plants . To adapt Cas9 for live cell imaging in plants, we introduced a point mutation in the RuvC1 and HNH nuclease domains (D10A and H841A) in two Cas9 orthologues derived from *S. pyogenes* (Sp-dCas9) and *S. aureus* (Sa-dCas9), rendering Cas9 nuclease-deficient (Figure 7).

Figure 7 Structure of the CRISPR-dCas9 construct.
Transcription of Sp/Sa-dCas9 was initiated by the parsley ubiquitin 4 promoter and terminated by the pea 3A terminator. An SV40 NLS DNA sequence was used for nuclear localization of dCas9. Transcription of the sgRNA scaffold was initiated by the *Arabidopsis* ubiquitin 6 promoter.

To functionally test CRISPR-dCas9 live cell imaging in plants, we decided to use *Nicotiana benthamiana* as a model species due to its relative ease of transformation via

Agrobacterium tumefaciens and comparably large telomere tandem repeat arrays (Fajkus *et al*, 1995; Goodin *et al*, 2008). At first, we interrogated the specificity at which CRISPR-dCas9 is capable of labelling telomeres during interphase. Compared to the number of telomeres detected by FISH, which was used as a reference, CRISPR-dCas9 labelled 78% of all telomeres. Importantly, no distinct signals were observed when CRISPR-dCas9 was used without a telomere-specific single guide (sg) RNA. Although it is difficult to compare different studies due to various species and experimental systems being used (e.g. mammalian cells, different DNA-binding proteins), it seems to be reasonable to expect a telomere labelling efficiency of about 70 - 85% of both CRISPR-dCas9 and TALE approaches (Chen *et al*, 2013; Fujimoto *et al*, 2016). Importantly, CRISPR-dCas9 labelling correlates well with FISH regarding fluorescence signal intensity and size ($\rho = 0.84$, $r^2 = 0.7$). Furthermore, we tested two different Cas9 orthologues and compared their efficiency. Both Cas9 orthologues from *S. pyogenes* and *S. aureus* were labelled with different fluorescence proteins and used simultaneously, and showed almost complete co-localization at the telomeres. This means both variants are highly efficient and can potentially be used to visualize different DNA sequences at the same time.

The ability to visualize protein-DNA interactions in living cells is intriguing as it might allow us to directly observe the dynamics between DNA-binding proteins, such as transcription factors, and their target sequence in the context of plant development or environmental stress. We therefore explored this possibility by simultaneously labelling telomeric DNA through CRISPR-dCas9 and the telomeric repeat binding protein 1 (TRB1) which is located at telomeres and interacts with telomerase reverse transcriptase (TERT, (Dvorácková *et al*, 2010; Schrumpfová *et al*, 2014)). Interestingly, telomeres in yeast, ciliates, mammals and plants may form 3' overhangs. However, in plants, blunt-ended telomeres and telomeres with 3' overhangs may appear in the same cell, and only telomeres exhibiting 3' overhangs are bound by TRB1. We found that 87.6 % of telomeres visualized by CRISPR-dCas9 were also bound by TRB1, which indicates that most telomeres in *N. benthamiana* form 3' overhangs and only a small proportion of blunt-ended telomeres are present during interphase.

We then attempted to explore the possibility of using CRISPR-dCas9 for live imaging of telomeres in *N. benthamiana* interphase nuclei. To better map telomeres within the nucleus, we additionally visualized the nuclear envelope via co-expression of the nuclear egress protein of the human cytomegalovirus pUL50 fused to GFP (Lamm *et al*, 2016).
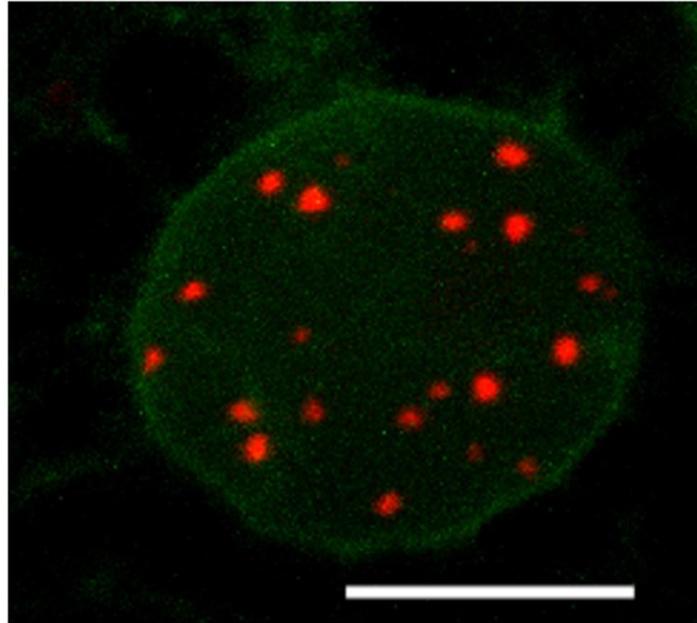
Figure 8 CRISPR-dCas9 live cell imaging
Live interphase nucleus of *N. benthamiana* showing telomeres (Sp-dCas9-mRuby, red) and nuclear envelope (pUL50-GFP, green). Standard bar equals 10 µm.

By conduction fluorescence recovery after photobleaching (FRAP) experiments, we first found that CRISPR-dCas9 is stably associated with its target sequence, which is in agreement with previous reports (Ma *et al*, 2016; Qin *et al*, 2017). We then sought to investigate telomere dynamics during interphase. As mentioned above (see 1.1.2.), chromatin dynamics, especially heterochromatic regions such as telomeres, can be related to transcriptional activation, replication, and repair. Although differentiated *N. benthamiana* leaf cells do not undergo cell division anymore, they might still respond to environmental stimuli and undergo transcriptional activation or silencing. By tracking single telomere clusters over time, we observed confined diffusion of telomeres as well as longe-range movements. Spatially confined telomere movements became evident by the almost linear ascent of the mean square displacement (MSD). However, we also

found a high standard deviation of the MSD which points towards strong variation between individual telomere clusters. These variations became evident as long-range movements of telomere clusters of up to 2 µm which amounts to 13% of the average nuclear diameter. A spatially confined diffusion of telomeres was previously observed in *A. thaliana* using fluorescent TALEs, which supports our observation (Fujimoto *et al*, 2016). However, such additional long-range telomere movements in live plant cells were not reported before and might be related to transcription of telomeric tandem repeats (Koo *et al*, 2016), telomerase activity (Procházková Schrumpfová *et al*, 2016), or positional silencing by telomeres (Gottschling *et al*, 1990; Nimmo *et al*, 1994; Cryderman *et al*, 1999). Based on the technical reliability of CRISPR-dCas9 imaging, I propose that these methodologies can be used in future work aiming to investigate spatio-temporal protein-DNA interactions under varying environmental conditions to elucidate how structural chromatin changes are translated into functional responses. As an example, our preliminary work on transferring the CRISPR-dCas9 imaging technique to *A. thaliana* in a stable manner appeared to be promising regarding the visualization of centromeric DNA (Figure 9). In future studies, this will be combined with fluorescent-tagged centromeric histone H3 (CENH3-GFP) to investigate dynamic associations between centromeric DNA and a functionally essential centromeric protein in living cells.
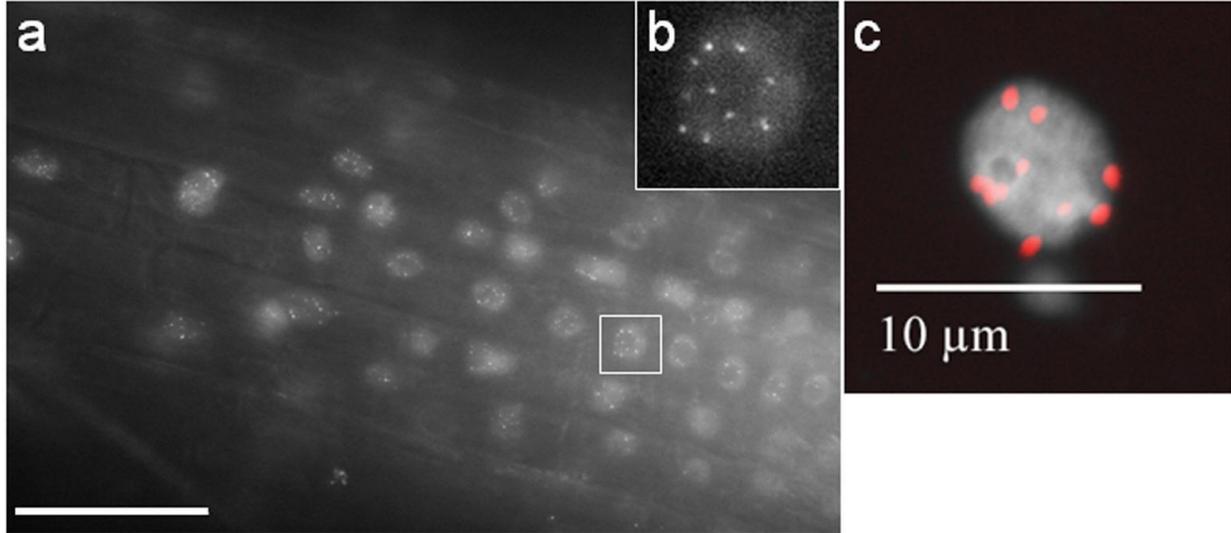
Figure 9 Application of CRISPR-imaging in *A. thaliana*.
(a) CRISPR-dCas9 derived from *S. aureus* was fused with three copies of far red fluorescent protein (TurboFP635) and directed to *A. thaliana* centromeric DNA via an sgRNA specific to the centromeric satellite repeat pAL1 (sgRNA protospacer sequence: 5'-CACTACTTAGGCTTTTAAGA-3'). Centromeric fluorescence signals within nuclei-like circular structures were observed in a living root of *A. thaliana*. Standard bar equals 50 µm. **(b)** Inset shows 10 bright fluorescence signals which resemble the expected 10 centromere signals in a diploid root cell. **(c)** Example of FISH used to visualize centromeres in an *A. thaliana* interphase nucleus.

4. Conclusions

The main conclusions of my study are that novel single-cell analysis methodologies for plant research were developed, their technical robustness was demonstrated by experimental evidence, and these methodologies were used to gain new insights into segregation and dynamics of defined chromosomal regions during meiosis and interphase.

First, we demonstrated the feasibility of flow sorting individual pollen nuclei followed by whole-genome-amplification via two different methods (MDA and PicoPLEX) to investigate the recombination landscape of barley at the molecular level in an unbiased way. Sequencing of individual pollen nuclei revealed that segregation distortion is almost absent, whereas it is detectable to a large extent in barley doubled haploid plants likely caused by selection during doubled haploid production. In addition, we presented meiotic recombination measurements which support the existence of interference sensitive (class I) and less sensitive (class II) crossovers in barley.

Second, we established live cell imaging of telomeres by repurposing the recently discovered CRISPR-Cas9 system. We demonstrated that the feasibility of CRISPR-Cas9 to visualize telomeres in the model plant *N. benthamiana* is comparable to other systems such as mammalian cells. Finally, using this method revealed how variable telomere clusters are in terms of their movements during interphase. We concluded that these dynamics can be a prerequisite for or result of transcriptional activation or silencing in response to environmental stimuli.

Together these methodologies pave the way for future research in the field of basic and applied plant sciences.

5. Summary

Single-cell analysis via whole-genome sequencing or live cell imaging is an exciting field in which many unknown factors of fundamental cellular processes remain to be discovered. New methodologies are constantly being developed for many model organisms belonging to different species. The current study deals with advancing two different single-cell analysis methodologies in plants.

First, single-cell whole-genome sequencing was established in barley, a large genome crop species important for human nutrition, for the purpose of directly measuring meiotic recombination frequency at the DNA sequence level. The aim of this part was to address a major issue of meiosis research, namely the low potential of sample throughput and limited resolution via microscopy-based methods. While single-cell genome sequencing achieves very high resolution (1 Mbp) that outperforms conventional microscopy, it also holds the potential to enable the analysis of thousands of samples since pollen nuclei can be sorted into individual reactions tubes via flow cytometry. However, the costs of library preparation for sequencing remain to be reduced to make this a widely applicable, high-throughput approach. Applying this methodology in barley pollen enabled us to provide supporting evidence for two meiotic crossover classes in this species, and revealed that segregation distortion in doubled haploid plants is not directly caused by meiosis but rather explained by selection during early plant development.

In the second part of this study, the recently discovered CRISPR-Cas9 system was repurposed for live cell imaging of telomeric DNA repeats in *N. benthamiana*. Visualizing specific genomic regions in the nuclei of living cells might enable us to improve our understanding of the spatio-temporal dynamics therein. Chromatin dynamics during interphase are associated with transcriptional activation or silencing of genes in response to environmental stimuli or developmental processes. It is therefore of great interest to develop new methodologies for their visualization. The aim of this part was to establish nuclease-deficient CRISPR-dCas9 for live cell imaging in plants and demonstrate its feasibility to observe spatio-temporal dynamics. We were able to demonstrate robust labelling of telomeres by CRISPR-dCas9 in living interphase nuclei of *N. benthamiana*, show its potential for protein-DNA interaction studies by simultaneously labelling a telomere repeat binding protein (TRB1), and pave the way for

future labelling of multiple genomic regions by demonstrating the feasibility of two different Cas9 orthologues. We were able to reveal great variability in the spatio-temporal dynamics of telomere clusters by tracking their movements over time in living leaf cells. The underlying mechanisms might be associated with transcriptional silencing through telomeric heterochromatin and may contribute to responses to environmental stimuli or developmental processes.

6. Zusammenfassung

Die Analyse einzelner Zellen mittels Genomsequenzierung oder der Mikroskopie von lebenden Zellen ist ein sich schnell entwickelndes Forschungsgebiet. Neue Methoden werden in diesem Gebiet kontinuierlich für zahlreiche Modellorganismen weiterentwickelt. Die vorliegende Studie beschäftigt sich mit der Etablierung von zwei unterschiedlichen Methoden der Einzelzellanalyse in Pflanzen.

Um auf direkte Weise meiotische Rekombinationsvorgänge auf DNA-Sequenzebene zu messen, wurde im ersten Teil dieser Studie die Genomsequenzierung einzelner Pollenzellkerne in der Kulturpflanze Gerste etabliert. Ziel war es, eine der technischen Schwierigkeiten in der Meioseforschung, nämlich das geringe Durchsatzpotenzial und die begrenzte Auflösung konventioneller Mikroskopiemethoden, zu adressieren. Die entwickelte Einzelpollenanalyse ermöglicht eine Auflösung von ~1 Mbp, welche der konventionellen Mikroskopie überlegen ist. Weiterhin besteht die Möglichkeit einzelne Zellen mittels Durchflusszytometrie zu sortieren und somit tausende von Stichproben zu generieren. Ein Nachteil liegt allerdings in den hohen Kosten der Sequenzierung, welche noch reduziert werden müssen, um eine breite Anwendung der Methode zu ermöglichen. Die Nutzung dieser Methodik hat zu folgenden Erkenntnissen geführt. So konnte die Existenz von zwei unterschiedlichen Klassen von meiotischen Rekombinationsvorgängen mit experimentellen Daten unterstützt werden. Weiterhin konnte gezeigt werden, dass Abweichungen von Mendel'schen Segregationsverhältnissen in doppel-haploider Gerste nicht direkt durch meiotische Prozesse verursacht wird, da normale Segregationsverhältnisse in Pollen gemessen wurden, sondern eher durch Selektionsdruck während der frühen Pflanzenentwicklung entstehen.

Im zweiten Teil dieser Studie wurde das kürzlich entdeckte CRISPR-Cas9 System für die Visualisierung von Telomer DNA Sequenzen in lebenden *N. benthamiana* Zellen eingesetzt. Die Visualisierung von definierten genomischen Bereichen dient der Erforschung von räumlich-zeitlichen Dynamiken im Zellkern. Solche Chromatindynamiken, welche während der Interphase auftreten, sind unter anderem mit der transkriptionellen Aktivierung oder Inaktivierung von Genen in Bezug auf Umwelteinflüsse oder Entwicklungsprozesse assoziiert. Es ist daher von Interesse neue

Methoden zur Visualisierung solcher Dynamiken zu entwickeln. Ziel war es, eine Nuklease-inaktive CRISPR-dCas9 Variante in Pflanzen zu etablieren und die Eignung dieser Methodik zur Visualisierung räumlich-zeitlicher Dynamiken zu demonstrieren. Es wurde gezeigt, dass Telomer DNA Sequenzen zuverlässig mittels CRISPR-dCas9 in lebenden Interphasezellkernen sichtbar gemacht werden können. Weiterhin wurde das Potenzial dieser Methodik zur Visualisierung von Protein-DNA Interaktionen demonstriert. Dazu wurde zusätzlich ein telomerspezifisches Protein (TRB1) mittels GFP markiert. Zuletzt wurde durch die simultane Anwendung von zwei unterschiedlichen Cas9 Orthologen das Potenzial dieser Methodik zur Visualisierung unterschiedlicher DNA Sequenzen gezeigt. Die Anwendung dieser Methodik ermöglichte es große Variabilität in der räumlich-zeitlichen Dynamik von Telomerclustern aufzuzeigen. Die zugrundeliegenden Mechanismen könnten mit der transkriptionellen Inaktivierung durch Telomer-spezifisches Heterochromatin assoziiert sein und in Bezug zu Reaktionen auf Umwelteinflüsse oder Entwicklungsprozesse stehen.

7. References

Aliyeva-Schnorr L, Beier S, Karafiátová M, Schmutzer T, Scholz U, Doležel J, Stein N & Houben A (2015) Cytogenetic mapping with centromeric bacterial artificial chromosomes contigs shows that this recombination-poor region comprises more than half of barley chromosome 3H. *Plant J.* **84:** 385–394

Anderson LK, Doyle GG, Brigham B, Carter J, Hooker KD, Lai A, Rice M & Stack SM (2003) High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. *Genetics* **165:** 849–865

Anton T, Bultmann S, Leonhardt H & Markaki Y (2014) Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. *Nucleus* **5:** 163–172

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408:** 796–815

Armstrong SJ, Franklin FC & Jones GH (2001) Nucleolus-associated telomere clustering and pairing precede meiotic chromosome synapsis in *Arabidopsis thaliana*. *J. Cell Sci.* **114:** 4207–4217

Baker K, Bayer M, Cook N, Dreißig S, Dhillon T, Russell J, Hedley PE, Morris J, Ramsay L, Colas I, Waugh R, Steffenson B, Milne I, Stephen G, Marshall D & Flavell AJ (2014) The low-recombining pericentromeric region of barley restricts gene diversity and evolution but not gene expression. *Plant J.* **79:** 981–992

Baker K, Dhillon T, Colas I, Cook N, Milne I, Milne L, Bayer M & Flavell AJ (2015) Chromatin state analysis of the barley epigenome reveals a higher-order structure defined by H3K27me1 and H3K27me3 abundance. *Plant J.* **84:** 111–124

Barcellona ML, Cardiel G & Gratton E (1990) Time-resolved fluorescence of DAPI in solution and bound to polydeoxynucleotides. *Biochem. Biophys. Res. Commun.* **170:** 270–280

Beerli RR & Barbas CF 3rd (2002) Engineering polydactyl zinc-finger transcription factors. *Nat. Biotechnol.* **20:** 135–141

Bélanger S, Clermont I, Esteves P & Belzile F (2016a) Extent and overlap of segregation

distortion regions in 12 barley crosses determined via a Pool-GBS approach. *Theor. Appl. Genet.* **129:** 1393–1404

Bélanger S, Esteves P, Clermont I, Jean M & Belzile F (2016b) Genotyping-by-Sequencing on pooled samples and its use in measuring segregation bias during the course of androgenesis in barley. *Plant Genome* **9:** Available at: http://dx.doi.org/10.3835/plantgenome2014.10.0073

Berchowitz LE & Copenhaver GP (2008) Fluorescent *Arabidopsis* tetrads: a visual assay for quickly developing large crossover and crossover interference data sets. *Nat. Protoc.* **3:** 41–50

Berchowitz LE & Copenhaver GP (2010) Genetic interference: don't stand so close to me. *Curr. Genomics* **11:** 91–102

Berchowitz LE, Francis KE, Bey AL & Copenhaver GP (2007) The role of AtMUS81 in interference-insensitive crossovers in *A. thaliana*. *PLoS Genet.* **3:** e132

Bergerat A, de Massy B, Gadelle D, Varoutas PC, Nicolas A & Forterre P (1997) An atypical topoisomerase II from *Archaea* with implications for meiotic recombination. *Nature* **386:** 414–417

Beumer K, Bhattacharyya G, Bibikova M, Trautman JK & Carroll D (2006) Efficient gene targeting in *Drosophila* with zinc-finger nucleases. *Genetics* **172:** 2391–2403

Bibikova M, Beumer K, Trautman JK & Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science* **300:** 764

Blainey PC, Mosier AC, Potanina A, Francis CA & Quake SR (2011) Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis. *PLoS One* **6:** e16626

Blainey PC & Quake SR (2014) Dissecting genomic diversity, one cell at a time. *Nat. Methods* **11:** 19–21

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A & Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326:** 1509–1512

Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, Imakaev M, Mirny LA, Wu C-T & Zhuang X (2016) Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529:** 418–422

Börner GV, Kleckner N & Hunter N (2004) Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis. *Cell* **117:** 29–45

de Bourcy CFA, De Vlaminck I, Kanbar JN, Wang J, Gawad C & Quake SR (2014) A quantitative comparison of single-cell whole genome amplification methods. *PLoS One* **9:** e105585

Cápal P, Blavet N, Vrána J, Kubaláková M & Doležel J (2015) Multiple displacement amplification of the DNA from single flow–sorted plant chromosome. *Plant J.* **84:** 838–844

Cápal P, Endo TR, Vrána J, Kubaláková M, Karafiátová M, Komínková E, Mora-Ramírez I, Weschke W & Doležel J (2016) The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat. *Plant Methods* **12:** 24

Chelysheva L, Diallo S, Vezon D, Gendrot G, Vrielynck N, Belcram K, Rocques N, Márquez-Lema A, Bhatt AM, Horlow C, Mercier R, Mézard C & Grelon M (2005) AtREC8 and AtSCC3 are essential to the monopolar orientation of the kinetochores during meiosis. *J. Cell Sci.* **118:** 4621–4632

Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, Park J, Blackburn EH, Weissman JS, Qi LS & Huang B (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155:** 1479–1491

Chen B, Hu J, Almeida R, Liu H, Balakrishnan S, Covill-Cooke C, Lim WA & Huang B (2016) Expanding the CRISPR imaging toolset with Staphylococcus aureus Cas9 for simultaneous imaging of multiple genomic loci. *Nucleic Acids Res.* **44:** e75

Chen P-H, Pan Y-B & Chen R-K (2008) High-throughput procedure for single pollen grain collection and polymerase chain reaction in plants. *J. Integr. Plant Biol.* **50:** 375–383

Choi K & Henderson IR (2015) Meiotic recombination hotspots - a comparative view. *Plant J.* **83:** 52–61

Choi K, Zhao X, Kelly KA, Venn O, Higgins JD, Yelina NE, Hardcastle TJ, Ziolkowski PA,

Copenhaver GP, Franklin FCH, McVean G & Henderson IR (2013) *Arabidopsis* meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nat. Genet.* **45:** 1327–1336

Choo Y, Sánchez-García I & Klug A (1994) In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence. *Nature* **372:** 642–645

Copenhaver GP, Keith KC & Preuss D (2000) Tetrad analysis in higher plants. A budding technology. *Plant Physiol.* **124:** 7–16

Couteau F, Belzile F, Horlow C, Grandjean O, Vezon D & Doutriaux MP (1999) Random chromosome segregation without meiotic arrest in both male and female meiocytes of a dmc1 mutant of *Arabidopsis*. *Plant Cell* **11:** 1623–1634

Cremer T & Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* **2:** 292–301

Cremer T & Cremer M (2010) Chromosome territories. *Cold Spring Harb. Perspect. Biol.* **2:** a003889

Cryderman DE, Morris EJ, Biessmann H, Elgin SC & Wallrath LL (1999) Silencing at *Drosophila* telomeres: nuclear organization and chromatin structure play critical roles. *EMBO J.* **18:** 3724–3735

Da Ines O, Abe K, Goubely C, Gallego ME & White CI (2012) Differing requirements for RAD51 and DMC1 in meiotic pairing of centromeres and chromosome arms in *Arabidopsis thaliana*. *PLoS Genet.* **8:** e1002636

Dean FB, Nelson JR, Giesler TL & Lasken RS (2001) Rapid amplification of plasmid and phage DNA using phi29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* **11:** 1095–1099

De Storme N & Geelen D (2011) The *Arabidopsis* mutant jason produces unreduced first division restitution male gametes through a parallel/fused spindle mechanism in meiosis II. *Plant Physiol.* **155:** 1403–1415

Dewitte W & Murray JAH (2003) The plant cell cycle. *Annu. Rev. Plant Biol.* **54:** 235–264

Dohm JC, Minoche AE, Holtgräwe D, Capella-Gutiérrez S, Zakrzewski F, Tafer H, Rupp O,

Sörensen TR, Stracke R, Reinhardt R, Goesmann A, Kraft T, Schulz B, Stadler PF, Schmidt T, Gabaldón T, Lehrach H, Weisshaar B & Himmelbauer H (2014) The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). *Nature* **505:** 546–549

Dong F & Jiang J (1998) Non-Rabl patterns of centromere and telomere distribution in the interphase nuclei of plant cells. *Chromosome Research* **6:** 551–558

Dreissig S, Fuchs J, Cápal P, Kettles N, Byrne E & Houben A (2015) Measuring meiotic crossovers via multi-locus genotyping of single pollen grains in barley. *PLoS One* **10:** e0137677

Dreissig S, Fuchs J, Himmelbach A, Mascher M & Houben A (2017) Sequencing of single pollen nuclei reveals meiotic recombination events at megabase resolution and circumvents segregation distortion caused by postmeiotic processes. *Front. Plant Sci.* Available at: http://dx.doi.org/10.3389/fpls.2017.01620

Drouaud J, Khademian H, Giraut L, Zanni V, Bellalou S, Henderson IR, Falque M & Mézard C (2013) Contrasted patterns of crossover and non-crossover at *Arabidopsis thaliana* meiotic recombination hotspots. *PLoS Genet.* **9:** e1003922

Drouaud J & Mézard C (2011) Characterization of meiotic crossovers in pollen from *Arabidopsis thaliana*. *Methods Mol. Biol.* **745:** 223–249

Dubin MJ, Zhang P, Meng D, Remigereau M-S, Osborne EJ, Paolo Casale F, Drewe P, Kahles A, Jean G, Vilhjálmsson B, Jagoda J, Irez S, Voronin V, Song Q, Long Q, Rätsch G, Stegle O, Clark RM & Nordborg M (2015) DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *Elife* **4:** e05255

Dvorácková M, Rossignol P, Shaw PJ, Koroleva OA, Doonan JH & Fajkus J (2010) AtTRB1, a telomeric DNA-binding protein from *Arabidopsis*, is concentrated in the nucleolus and shows highly dynamic association with chromatin. *Plant J.* **61:** 637–649

Fajkus J, Kovarík A, Královics R & Bezděk M (1995) Organization of telomeric and subtelomeric chromatin in the higher plant *Nicotiana tabacum*. *Mol. Gen. Genet.* **247:** 633–638

Fauser F, Schiml S & Puchta H (2014) Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J.* **79:** 348–359

Feng C-M, Qiu Y, Van Buskirk EK, Yang EJ & Chen M (2014a) Light-regulated gene repositioning in *Arabidopsis. Nat. Commun.* **5:** 3027

Feng S, Cokus SJ, Schubert V, Zhai J, Pellegrini M & Jacobsen SE (2014b) Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in *Arabidopsis. Mol. Cell* **55:** 694–707

Ferdous M, Higgins JD, Osman K, Lambing C, Roitinger E, Mechtler K, Armstrong SJ, Perry R, Pradillo M, Cuñado N & Franklin FCH (2012) Inter-homolog crossing-over and synapsis in *Arabidopsis* meiosis are dependent on the chromosome axis protein AtASY3. *PLoS Genet.* **8:** e1002507

Fischer A, Hofmann I, Naumann K & Reuter G (2006) Heterochromatin proteins and the control of heterochromatic gene silencing in *Arabidopsis. J. Plant Physiol.* **163:** 358–368

Francis KE, Lam SY, Harrison BD, Bey AL, Berchowitz LE & Copenhaver GP (2007) Pollen tetrad-based visual assay for meiotic recombination in *Arabidopsis. Proceedings of the National Academy of Sciences* **104:** 3913–3918

Fujimoto S, Sugano SS, Kuwata K, Osakabe K & Matsunaga S (2016) Visualization of specific repetitive genomic sequences with fluorescent TALEs in *Arabidopsis thaliana. J. Exp. Bot.* **67:** 6101–6110

Funabiki H, Hagan I, Uzawa S & Yanagida M (1993) Cell cycle-dependent specific positioning and clustering of centromeres and telomeres in fission yeast. *J. Cell Biol.* **121:** 961–976

Gale MD, Rees H & Others (1970) Genes controlling chiasma frequency in *Hordeum. Heredity* **25:** 393–410

Gawad C, Koh W & Quake SR (2016) Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* **17:** 175–188

Gilson E, Laroche T & Gasser SM (1993) Telomeres and the functional architecture of the nucleus. *Trends Cell Biol.* **3:** 128–134

Giraut L, Falque M, Drouaud J, Pereira L, Martin OC & Mézard C (2011) Genome-wide crossover distribution in *Arabidopsis thaliana* meiosis reveals sex-specific patterns along chromosomes. *PLoS Genet.* **7:** e1002354

Gole J, Gore A, Richards A, Chiu Y-J, Fung H-L, Bushman D, Chiang H-I, Chun J, Lo Y-H & Zhang K (2013) Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nat. Biotechnol.* **31:** 1126–1132

Goodin MM, Zaitlin D, Naidu RA & Lommel SA (2008) *Nicotiana benthamiana*: Its history and future as a model for plant–pathogen interactions. *Mol. Plant. Microbe. Interact.* **21:** 1015–1026

Goss JA (1968) Development, physiology, and biochemistry of corn and wheat pollen. *Bot. Rev.* **34:** 333–359

Gottschling DE, Aparicio OM, Billington BL & Zakian VA (1990) Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* **63:** 751–762

Grelon M, Vezon D, Gendrot G & Pelletier G (2001) AtSPO11-1 is necessary for efficient meiotic recombination in plants. *EMBO J.* **20:** 589–600

Grob S, Schmid MW & Grossniklaus U (2014) Hi-C analysis in *Arabidopsis* identifies the KNOT, a structure with similarities to the flamenco locus of *Drosophila*. *Mol. Cell* **55:** 678–693

Habu Y, Ando T, Ito S, Nagaki K, Kishimoto N, Taguchi-Shiobara F, Numa H, Yamaguchi K, Shigenobu S, Murata M, Meshi T & Yano M (2015) Epigenomic modification in rice controls meiotic recombination and segregation distortion. *Mol. Breed.* **35:** 103

Haraguchi T (2002) Live cell imaging: approaches for studying protein dynamics in living cells. *Cell Struct. Funct.* **27:** 333–334

Hartung F, Wurz-Wildersinn R, Fuchs J, Schubert I, Suer S & Puchta H (2007) The catalytically active tyrosine residues of both SPO11-1 and SPO11-2 are required for meiotic double-strand break induction in *Arabidopsis*. *Plant Cell* **19:** 3090–3099

Hess M, Sczyrba A, Egan R, Kim T-W, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T, Mackie RI, Pennacchio LA, Tringe SG, Visel A, Woyke T, Wang Z & Rubin EM (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* **331:** 463–467

Higgins JD, Perry RM, Barakate A, Ramsay L, Waugh R, Halpin C, Armstrong SJ & Franklin FCH (2012) Spatiotemporal asymmetry of the meiotic program underlies the predominantly

distal distribution of meiotic crossovers in barley. *Plant Cell* **24:** 4096–4109

Houben A, Kynast RG, Heim U, Hermann H, Jones RN & Forster JW (1996) Molecular cytogenetic characterisation of the terminal heterochromatic segment of the B-chromosome of rye (*Secale cereale*). *Chromosoma* **105:** 97–103

International Barley Genome Sequencing Consortium, Mayer KFX, Waugh R, Brown JWS, Schulman A, Langridge P, Platzer M, Fincher GB, Muehlbauer GJ, Sato K, Close TJ, Wise RP & Stein N (2012) A physical, genetic and functional sequence assembly of the barley genome. *Nature* **491:** 711–716

International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463:** 763–768

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature* **436:** 793–800

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA & Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337:** 816–821

Kalisky T, Blainey P & Quake SR (2011) Genomic analysis at the single-cell level. *Annu. Rev. Genet.* **45:** 431–445

Kato N & Lam E (2001) Detection of chromosomes tagged with green fluorescent protein in live *Arabidopsis thaliana* plants. *Genome Biol.* **2:** RESEARCH0045

Kawakatsu T, Huang S-SC, Jupe F, Sasaki E, Schmitz RJ, Urich MA, Castanon R, Nery JR, Barragan C, He Y, Chen H, Dubin M, Lee C-R, Wang C, Bemm F, Becker C, O'Neil R, O'Malley RC, Quarless DX, 1001 Genomes Consortium, et al (2016) Epigenomic diversity in a global collection of arabidopsis thaliana accessions. *Cell* **166:** 492–505

Keeney S, Giroux CN & Kleckner N (1997) Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88:** 375–384

Khademian H, Giraut L, Drouaud J & Mézard C (2013) Characterization of meiotic non-crossover molecules from *Arabidopsis thaliana* pollen. *Methods Mol. Biol.* **990:** 177–190

Klug A (1993) Co-chairman's remarks: protein designs for the specific recognition of DNA. *Gene* **135:** 83–92

Koo D-H, Zhao H & Jiang J (2016) Chromatin-associated transcripts of tandemly repetitive DNA sequences revealed by RNA-FISH. *Chromosome Res.* **24:** 467–480

Kozubek S, Lukásová E, Amrichová J, Kozubek M, Lisková A & Slotová J (2000) Influence of cell fixation on chromatin topography. *Anal. Biochem.* **282:** 29–38

Kubista M, Aakerman B & Norden B (1987) Characterization of interaction between DNA and 4',6-diamidino-2-phenylindole by optical spectroscopy. *Biochemistry* **26:** 4545–4553

Lam E, Kato N & Watanabe K (2004) Visualizing chromosome structure/organization. *Annu. Rev. Plant Biol.* **55:** 537–554

Lamm CE, Link K, Wagner S, Milbradt J, Marschall M & Sonnewald U (2016) Human cytomegalovirus nuclear egress proteins ectopically expressed in the heterologous environment of plant cells are strictly targeted to the nuclear envelope. *Viruses* **8:** 73

Langmore JP (2002) Rubicon Genomics, Inc. *Pharmacogenomics* **3:** 557–560

Leung ML, Wang Y, Waters J & Navin NE (2015) SNES: single nucleus exome sequencing. *Genome Biol.* **16:** 55

Lindhout BI, Fransz P, Tessadori F, Meckel T, Hooykaas PJJ & van der Zaal BJ (2007) Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. *Nucleic Acids Res.* **35:** e107

Li X, Li L & Yan J (2015) Dissecting meiotic recombination based on tetrad analysis by single-microspore sequencing in maize. *Nat. Commun.* **6:** 6648

Llères D, James J, Swift S, Norman DG & Lamond AI (2009) Quantitative analysis of chromatin compaction in living cells using FLIM–FRET. *J. Cell Biol.* **187:** 481–496

Lu P, Han X, Qi J, Yang J, Wijeratne AJ, Li T & Ma H (2012) Analysis of *Arabidopsis* genome-wide variations before and after meiosis and meiotic recombination by resequencing Landsberg erecta and all four products of a single meiosis. *Genome Res.* **22:** 508–518

Lynch M (2010) Evolution of the mutation rate. *Trends Genet.* **26:** 345–352

Lysak MA, Fransz PF, Ali HB & Schubert I (2001) Chromosome painting in *Arabidopsis thaliana*. *Plant J.* **28:** 689–697

Ma H, Naseri A, Reyes-Gutierrez P, Wolfe SA, Zhang S & Pederson T (2015) Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **112:** 3002–3007

Ma H, Reyes-Gutierrez P & Pederson T (2013) Visualization of repetitive DNA sequences in human chromosomes with transcription activator-like effectors. *Proc. Natl. Acad. Sci. U. S. A.* **110:** 21048–21053

Ma H, Tu L-C, Naseri A, Huisman M, Zhang S, Grunwald D & Pederson T (2016) CRISPR-Cas9 nuclear dynamics and target recognition in living cells. *J. Cell Biol.* **214:** 529–537

Manninen OM (2000) Associations between anther-culture response and molecular markers on chromosomes 2H, 3H and 4H of barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **100:** 57–62

Manuelidis L (1984) Different central nervous system cell types display distinct and nonrandom arrangements of satellite DNA sequences. *Proc. Natl. Acad. Sci. U. S. A.* **81:** 3123–3127

Marcy Y, Ouverney C, Bik EM, Lösekann T, Ivanova N, Martin HG, Szeto E, Platt D, Hugenholtz P, Relman DA & Quake SR (2007) Dissecting biological 'dark matter' with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proceedings of the National Academy of Sciences* **104:** 11889–11894

Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V, Dockter C, Hedley PE, Russell J, Bayer M, Ramsay L, Liu H, Haberer G, Zhang X-Q, Zhang Q, Barrero RA, Li L, Taudien S, Groth M, et al (2017) A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544:** 427–433

Matsunaga S, Schütze K, Donnison IS, Grant SR, Kuroiwa T & Kawano S (1999) Single pollen typing combined with laser-mediated manipulation. *Plant J.* **20:** 371–378

Mayr E (1982) The growth of biological thought: diversity, evolution, and inheritance. Harvard University Press

Melamed-Bessudo C & Levy AA (2012) Deficiency in DNA methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* **109:** E981–8

Mercier R, Mézard C, Jenczewski E, Macaisne N & Grelon M (2015) The molecular biology of

meiosis in plants. *Annu. Rev. Plant Biol.* **66:** 297–327

Miller J, McLachlan AD & Klug A (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. *EMBO J.* **4:** 1609–1614

Mirouze M, Lieberman-Lazarovich M, Aversano R, Bucher E, Nicolet J, Reinders J & Paszkowski J (2012) Loss of DNA methylation affects the recombination landscape in *Arabidopsis. Proc. Natl. Acad. Sci. U. S. A.* **109:** 5880–5885

Moghaddam A, Roudier F & Seifert M (2011) Additive inheritance of histone modifications in *Arabidopsis thaliana* intra-specific hybrids. *Plant J.* Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1365-313X.2011.04628.x/full

Nelles DA, Fang MY, O'Connell MR, Xu JL, Markmiller SJ, Doudna JA & Yeo GW (2016) Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell* **165:** 488–496

Nilsson N-O & Pelger S (1991) The relationship between natural variation in chiasma frequencies and recombination frequencies in barley. *Hereditas* **115:** 121–126

Nimmo ER, Cranston G & Allshire RC (1994) Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes. *EMBO J.* **13:** 3801–3811

O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M & Doudna JA (2014) Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **516:** 263–266

Osman K, Sanchez-Moran E, Mann SC, Jones GH & Franklin FCH (2009) Replication protein A (AtRPA1a) is required for class I crossover formation but is dispensable for meiotic DNA break repair. *EMBO J.* **28:** 394–404

Pabo CO, Peisach E & Grant RA (2001) Design and selection of novel Cys2His2 zinc finger proteins. *Annu. Rev. Biochem.* **70:** 313–340

Pacher M & Puchta H (2017) From classical mutagenesis to nuclease-based breeding--directing natural DNA repair for a natural end-product. *Plant J.* **90:** 819–833

Pavletich NP & Pabo CO (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252:** 809–817

Pecinka A, Dinh HQ, Baubec T, Rosa M, Lettner N & Mittelsten Scheid O (2010) Epigenetic

regulation of repetitive elements is attenuated by prolonged heat stress in *Arabidopsis*. *Plant Cell* **22:** 3118–3129

Petersen G, Johansen B & Seberg O (1996) PCR and sequencing from a single pollen grain. *Plant Mol. Biol.* **31:** 189–191

Phillips D, Jenkins G, Macaulay M, Nibau C, Wnetrzak J, Fallding D, Colas I, Oakey H, Waugh R & Ramsay L (2015) The effect of temperature on the male and female recombination landscape of barley. *New Phytol.* **208:** 421–429

Phillips D, Nibau C, Wnetrzak J & Jenkins G (2012) High resolution analysis of meiotic chromosome structure and behaviour in barley (Hordeum vulgare L.). *PLoS One* **7:** e39539

Phillips D, Wnetrzak J, Nibau C, Barakate A, Ramsay L, Wright F, Higgins JD, Perry RM & Jenkins G (2013) Quantitative high resolution mapping of HvMLH3 foci in barley pachytene nuclei reveals a strong distal bias and weak interference. *J. Exp. Bot.* **64:** 2139–2154

Piatek A, Ali Z, Baazim H, Li L, Abulfaraj A, Al-Shareef S, Aouida M & Mahfouz MM (2015) RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnol. J.* **13:** 578–589

Prieto P, Shaw P & Moore G (2004) Homologue recognition during meiosis is associated with a change in chromatin conformation. *Nat. Cell Biol.* **6:** 906–908

Procházková Schrumpfová P, Schořová Š & Fajkus J (2016) Telomere- and telomerase-associated proteins and their functions in the plant cell. *Front. Plant Sci.* **7:** 851

Qi J, Chen Y, Copenhaver GP & Ma H (2014) Detection of genomic variations and DNA polymorphisms and impact on analysis of meiotic recombination and genetic mapping. *Proc. Natl. Acad. Sci. U. S. A.* **111:** 10007–10012

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP & Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152:** 1173–1183

Qin P, Parlak M, Kuscu C, Bandaria J, Mir M, Szlachta K, Singh R, Darzacq X, Yildiz A & Adli M (2017) Live cell imaging of low- and non-repetitive chromosome loci using CRISPR-Cas9. *Nat. Commun.* **8:** 14725

Rabl C (1885) Über Zellteilung. *Morphologisches Jahrbuch* **10:** 214–330

Rawlins DJ & Shaw PJ (1990) Three-dimensional organization of ribosomal DNA in interphase nuclei of *Pisum sativum* by in situ hybridization and optical tomography. *Chromosoma* **99:** 143–151

Rodrigue S, Malmstrom RR, Berlin AM, Birren BW, Henn MR & Chisholm SW (2009) Whole genome amplification and de novo assembly of single bacterial cells. *PLoS One* **4:** e6864

Ropars J, Toro KS, Noel J, Pelin A, Charron P, Farinelli L, Marton T, Krüger M, Fuchs J, Brachmann A & Others (2016) Evidence for the sexual origin of heterokaryosis in arbuscular mycorrhizal fungi. *Nature microbiology* **1:** 16033

Saintenac C, Faure S, Remay A, Choulet F, Ravel C, Paux E, Balfourier F, Feuillet C & Sourdille P (2011) Variation in crossover rates across a 3-Mb contig of bread wheat (*Triticum aestivum*) reveals the presence of a meiotic recombination hotspot. *Chromosoma* **120:** 185–198

Säll T (1990) Genetic control of recombination in barley: Variation in linkage between marker genes. *Hereditas* **112:** 171–178

Säll T, Flink J & Bengtsson BO (1990) Genetic control of recombination in barley: Variation in recombination frequency measured with inversion heterozygotes. *Hereditas* **112:** 157–170

Salomé PA, Bomblies K, Fitz J, Laitinen RAE, Warthmann N, Yant L & Weigel D (2012) The recombination landscape in *Arabidopsis thaliana* F2 populations. *Heredity* **108:** 447–455

Schiml S, Fauser F & Puchta H (2014) The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in *Arabidopsis* resulting in heritable progeny. *Plant J.* **80:** 1139–1150

Schrumpfová PP, Vychodilová I, Dvořáčková M, Majerská J, Dokládal L, Schořová S & Fajkus J (2014) Telomere repeat binding proteins are functional components of *Arabidopsis* telomeres and interact with telomerase. *Plant J.* **77:** 770–781

Schubert I & Shaw P (2011) Organization and dynamics of plant interphase chromosomes. *Trends Plant Sci.* **16:** 273–281

Schubert V, Berr A & Meister A (2012) Interphase chromatin organisation in *Arabidopsis* nuclei:

constraints versus randomness. *Chromosoma* **121:** 369–387

Schubert V & Weisshart K (2015) Abundance and distribution of RNA polymerase II in *Arabidopsis* interphase nuclei. *J. Exp. Bot.* **66:** 1687–1698

Sekar RB & Periasamy A (2003) Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations. *J. Cell Biol.* **160:** 629–633

Shi J, Cui M, Yang L, Kim Y-J & Zhang D (2015) Genetic and biochemical mechanisms of pollen wall development. *Trends Plant Sci.* **20:** 741–753

Shilo S, Melamed-Bessudo C, Dorone Y, Barkai N & Levy AA (2015) DNA crossover motifs associated with epigenetic modifications delineate open chromatin regions in *Arabidopsis*. *Plant Cell* **27:** 2427–2436

Sidhu GK, Fang C, Olson MA, Falque M, Martin OC & Pawlowski WP (2015) Recombination patterns in maize reveal limits to crossover homeostasis. *Proc. Natl. Acad. Sci. U. S. A.* **112:** 15982–15987

Stacey NJ, Kuromori T, Azumi Y, Roberts G, Breuer C, Wada T, Maxwell A, Roberts K & Sugimoto-Shirasu K (2006) *Arabidopsis* SPO11-2 functions with SPO11-1 in meiotic recombination. *Plant J.* **48:** 206–216

Steinert J, Schiml S, Fauser F & Puchta H (2015) Highly efficient heritable plant genome engineering using Cas9 orthologues from *Streptococcus thermophilus* and *Staphylococcus aureus*. *Plant J.* **84:** 1295–1305

Sun Y, Ambrose JH, Haughey BS, Webster TD, Pierrie SN, Muñoz DF, Wellman EC, Cherian S, Lewis SM, Berchowitz LE & Copenhaver GP (2012) Deep genome-wide measurement of meiotic gene conversion using tetrad analysis in *Arabidopsis thaliana*. *PLoS Genet.* **8:** e1002968

Sybenga J (1966) The quantitative analysis of chromosome pairing and chiasma formation based on the relative frequencies of MI configurations. V. Interchange trisomics. *Genetica* **37:** 481–510

Telenius H, Carter NP, Bebb CE, Nordenskjöld M, Ponder BA & Tunnacliffe A (1992) Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single

degenerate primer. *Genomics* **13:** 718–725

Tiang C-L, He Y & Pawlowski WP (2012) Chromosome organization and dynamics during interphase, mitosis, and meiosis in plants. *Plant Physiol.* **158:** 26–34

Van't Hof J (1985) Control points within the cell cycle. *The cell division cycle in plants. Cambridge University Press, Cambridge***:** 1–13

Villeneuve AM & Hillers KJ (2001) Whence meiosis? *Cell* **106:** 647–650

Wang J, Fan HC, Behr B & Quake SR (2012) Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell* **150:** 402–412

Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, Chen K, Scheet P, Vattathil S, Liang H, Multani A, Zhang H, Zhao R, Michor F, Meric-Bernstam F & Navin NE (2014) Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512:** 155–160

Wijnker E, Velikkakam James G, Ding J, Becker F, Klasen JR, Rawat V, Rowan BA, de Jong DF, de Snoo CB, Zapata L, Huettel B, de Jong H, Ossowski S, Weigel D, Koornneef M, Keurentjes JJ & Schneeberger K (2013) The genomic landscape of meiotic crossovers and gene conversions in *Arabidopsis thaliana*. *Elife* **2:** e01426

Woyke T, Tighe D, Mavromatis K, Clum A, Copeland A, Schackwitz W, Lapidus A, Wu D, McCutcheon JP, McDonald BR, Moran NA, Bristow J & Cheng J-F (2010) One bacterial cell, one complete genome. *PLoS One* **5:** e10314

Woyke T, Xie G, Copeland A, González JM, Han C, Kiss H, Saw JH, Senin P, Yang C, Chatterji S, Cheng J-F, Eisen JA, Sieracki ME & Stepanauskas R (2009) Assembling the marine metagenome, one cell at a time. *PLoS One* **4:** e5299

Yang S, Yuan Y, Wang L, Li J, Wang W, Liu H, Chen J-Q, Hurst LD & Tian D (2012) Great majority of recombination events in *Arabidopsis* are gene conversion events. *Proc. Natl. Acad. Sci. U. S. A.* **109:** 20992–20997

Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, Miller N, Drouaud J, Grelon M, Copenhaver GP, Mezard C, Kelly KA & Henderson IR (2012) Epigenetic remodeling of meiotic crossover frequency in *Arabidopsis thaliana* DNA methyltransferase

mutants. *PLoS Genet.* **8:** e1002844

Yelina NE, Lambing C, Hardcastle TJ, Zhao X, Santos B & Henderson IR (2015) DNA
methylation epigenetically silences crossover hot spots and controls chromosomal domains
of meiotic recombination in *Arabidopsis. Genes Dev.* **29:** 2183–2202

Yelina NE, Ziolkowski PA, Miller N, Zhao X, Kelly KA, Muñoz DF, Mann DJ, Copenhaver GP &
Henderson IR (2013) High-throughput analysis of meiotic crossover frequency and
interference via flow cytometry of fluorescent pollen in *Arabidopsis thaliana. Nat. Protoc.* **8:**
2119–2134

Yuan K, Shermoen AW & O'Farrell PH (2014) Illuminating DNA replication during *Drosophila*
development using TALE-lights. *Curr. Biol.* **24:** R144–5

Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz
SE, Joung J, van der Oost J, Regev A, Koonin EV & Zhang F (2015) Cpf1 is a single RNA-
guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163:** 759–771

Zhang DY, Brandwein M, Hsuih T & Li HB (2001) Ramification amplification: a novel isothermal
DNA amplification method. *Mol. Diagn.* **6:** 141–150

Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, Li X, Pierick CJ,
Dobbs D, Peterson T, Joung JK & Voytas DF (2010) High frequency targeted mutagenesis
in *Arabidopsis thaliana* using zinc finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* **107:**
12028–12033

Zhang L, Cui X, Schmitt K, Hubert R, Navidi W & Arnheim N (1992) Whole genome amplification
from a single cell: implications for genetic analysis. *Proc. Natl. Acad. Sci. U. S. A.* **89:** 5847–
5851

Ziolkowski PA, Berchowitz LE, Lambing C, Yelina NE, Zhao X, Kelly KA, Choi K, Ziolkowska L,
June V, Sanchez-Moran E, Franklin C, Copenhaver GP & Henderson IR (2015)
Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover
remodelling via interference during *Arabidopsis* meiosis. *Elife* **4:** Available at:
http://dx.doi.org/10.7554/eLife.03708

Ziolkowski PA, Underwood CJ, Lambing C, Martinez-Garcia M, Lawrence EJ, Ziolkowska L,
Griffin C, Choi K, Franklin FCH, Martienssen RA & Henderson IR (2017) Natural variation

and dosage of the HEI10 meiotic E3 ligase control *Arabidopsis* crossover recombination. *Genes Dev.* **31:** 306–317

Zong C, Lu S, Chapman AR & Xie XS (2012) Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* **338:** 1622–1626

8. Abbreviations

bp	base pair
Cas9	CRISPR associated protein 9
CO	crossover
CRISPR	clustered regularly interspaced short palindromic repeats
DAPI	4',6-diamidino-2-phenylindole
dCas9	nuclease deficient Cas9
DH	double haploid
dHJ	double Holliday junction
DNA	deoxyribonucleic acid
DOP-PCR	degenerate oligonucleotide priming PCR
DSB	double-strand break
FACS	fluorescence activated cell sorting
FISH	fluorescence in situ hybridization
FRAP	fluorescence recovery after photobleaching
GFP	green fluorescence protein
KASP	competitive allele specific PCR
kbp	kilobase pair
µm	micrometer
MALBAC	multiple annealing and loop-based amplification cycles
Mbp	megabase pair
MDA	multiple displacement amplification
NCO	non-crossover
nm	nanometer
PCR	polymerase chain reaction
QTL	quantitative trait loci
RNA	ribonucleic acid
SDSA	synthesis-dependent strand annealing
SNP	single nucleotide polymorphism
TALE	transcription activator-like effector
WGA	whole-genome amplification
ZFP	zinc finger protein
ZMM	collective name for Zip1, Zip2, Zip3, Zip4, Mer3, Msh4, and Msh5 proteins

9. Acknowledgements

10. Selbstständigkeitserklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe und die den benutzten Werken wörtlich oder Inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

Datum / Date

the applicant

Unterschrift des Antragstellers / Signature of