



DeepTetrad: high-throughput image analysis of meiotic tetrads by deep learning in *Arabidopsis thaliana*

Eun-Cheon Lim¹, Jaeil Kim¹, Jihye Park¹, Eun-Jung Kim¹, Juhyun Kim¹, Yeong Mi Park¹, Hyun Seob Cho¹, Dohwan Byun¹, Ian R. Henderson², Gregory P. Copenhaver³ , Ildoo Hwang¹ and Kyuha Choi^{1,*} 

¹Department of Life Sciences, Pohang University of Science and Technology, Pohang, Gyeongbuk, Republic of Korea,

²Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK, and

³Department of Biology and the Integrative Program for Biological and Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Received 7 July 2019; revised 3 September 2019; accepted 9 September 2019; published online 19 September 2019.

*For correspondence (e-mail kyuha@postech.ac.kr).

SUMMARY

Meiotic crossovers facilitate chromosome segregation and create new combinations of alleles in gametes. Crossover frequency varies along chromosomes and crossover interference limits the coincidence of closely spaced crossovers. Crossovers can be measured by observing the inheritance of linked transgenes expressing different colors of fluorescent protein in *Arabidopsis* pollen tetrads. Here we establish DeepTetrad, a deep learning-based image recognition package for pollen tetrad analysis that enables high-throughput measurements of crossover frequency and interference in individual plants. DeepTetrad will accelerate the genetic dissection of mechanisms that control meiotic recombination.

Keywords: meiosis, crossover, interference, pollen tetrad analysis, fluorescent-tagged line, deep learning, technical advance.

INTRODUCTION

Meiosis consists of two consecutive nuclear divisions and produces four haploid gametes from a single diploid cell in sexually reproducing eukaryotes (Hunter, 2015). In *Arabidopsis* male meiosis, ~200–250 meiotic DNA double-stranded breaks (DSBs) are induced in the genome by a DNA topoisomerase VI-like complex to initiate meiotic recombination (Mercier *et al.*, 2015; Vrielynck *et al.*, 2016; Wang and Copenhaver, 2018). Of these DSBs, only ~8–11 are repaired as crossovers (COs) using a homologous chromosome (homolog). Thus, male meiosis in the *Arabidopsis* genome, which comprises five chromosomes, results in an average of ~1.8 crossovers between homologs. This low number suggests the existence of mechanisms that limit crossovers, a phenomenon that is observed in most eukaryotes (Mercier *et al.*, 2015). Meiotic DSB and CO frequencies are controlled by genetic and epigenetic factors and are non-randomly distributed along chromosomes, with higher levels around gene promoters and terminators and lower levels across the centromeres (Melamed-Bessudo and Levy, 2012; Yelina *et al.*, 2012; Wijnker *et al.*, 2013; Choi *et al.*, 2013; Shilo *et al.*, 2015; Choi and

Henderson, 2015; Melamed-Bessudo *et al.*, 2016; Choi *et al.*, 2018; Underwood *et al.*, 2018).

At least two pathways (type I and type II), contribute to CO formation (Mercier *et al.*, 2015; Wang and Copenhaver, 2018). The type I pathway leads to interfering COs that prevent the coincident occurrence of closely spaced CO on the same pair of chromosomes (Higgins *et al.*, 2004; Mercier *et al.*, 2005; Mercier *et al.*, 2015). In plants, interfering COs represent ~80–85% of total COs and are dependent on the ZMM proteins (ZIP4, MSH4, MSH5, MER3, HEI10, SHOC1, PTD, MLH1, MLH3). The remaining ~10–15% of COs are non-interfering and occur via the type II pathway (Berchowitz *et al.*, 2007). Non-interfering COs are resolved by the MUS81 endonuclease (Berchowitz *et al.*, 2007) and are restricted by anti-recombination factors such as FANCM, RECQ4A, RECQ4B, and FIGL1 (Crismani *et al.*, 2012; Girard *et al.*, 2015; Choi, 2017; Fernandes *et al.*, 2018; Serra *et al.*, 2018). Disruption of anti-recombination factors can increase the number of type II COs in plants, and this has the potential to create new combinations of desirable alleles that can improve crop varieties (Choi, 2017; Mieulet

et al., 2018). Therefore, high-throughput detection and understanding of CO frequency and interference have important implications for our understanding of the control of meiotic recombination as well as for breeding.

In Arabidopsis, CO frequency can be measured by pollen tetrad analysis using fluorescent-tagged lines (FTLs) in the *quartet1* (*qrt1*) background or by image analysis of fluorescent proteins in seed coats (Melamed-Bessudo *et al.*, 2005; Francis *et al.*, 2007; Berchowitz and Copenhaver, 2008; Yelina *et al.*, 2013; Ziolkowski *et al.*, 2015; van Tol *et al.*, 2018). Mutation of the *QRT1* gene encoding a pectin methylesterase results in the four-pollen products of male meiosis remaining attached to one another, allowing classical tetrad analysis. Each FTL contains a transgene that expresses eYFP (Y), dsRed (R) or eCFP (C) fluorescent proteins in mature pollen using the post-meiotic *LAT52* promoter (Twell *et al.*, 1990). Genetic intervals bounded by transgenes expressing different colors (e.g. *l1bc*, *l1fg*, *l2ab*, *l2fg*, *l3bc*, *CEN3*, *l5ab*; Figure S1) can be created by crossing FTLs. Scoring the segregation of two or three linked markers enables CO frequency and interference to be measured. For example, plants that are hemizygous for the three markers (*YRC*/+++) that define the *l1bc* interval produce 12 pollen tetrad classes (A–L) depending on the number of COs between *YR* and *RC* (Figure 1). The relative segregation of any two markers can be used to place pollen tetrads into one of the three categories used for classic tetrad analysis (Perkins, 1949): parental ditype (PD), tetra-type (T) and non-parental ditype (NPD) (Figure S2). Tetrad analysis enables the calculation of map distances between pairs of markers, and measurement of CO interference between adjacent intervals (Perkins, 1962).

Visual analysis of pollen tetrads is a powerful method for measuring genetic distance and crossover interference in Arabidopsis (Francis *et al.*, 2007; Berchowitz and Copenhaver, 2008; Crismani *et al.*, 2012; Girard *et al.*, 2015; Fernandes *et al.*, 2018). For example, manual analysis using fluorescence microscopy has been used to measure interference by comparing the map distances of two-color FTL intervals with and without COs in an adjacent interval (Berchowitz and Copenhaver, 2008). However, manually scoring large numbers of tetrads is laborious and time consuming. Alternatively, FTLs on the *qrt1*/+ or *QRT1* background can be analyzed by flow cytometry and allows rapid measurement of CO frequency and interference of ~10 000 single pollen grains per plant (Yelina *et al.*, 2013; Ziolkowski *et al.*, 2015; Ziolkowski *et al.*, 2017). Unfortunately, the flow cytometric method is unable to detect double crossovers within single intervals, requires high-purity pollen samples, and uses specialized equipment for three-color measurements.

Image software tools such as CellProfiler have been used for high-throughput analysis of single cells, including counting fluorescent seeds in Arabidopsis (Melamed-

Bessudo *et al.*, 2005; Carpenter *et al.*, 2006; Ziolkowski *et al.*, 2015; Ziolkowski *et al.*, 2017). Deep learning, a set of machine-learning approaches involving neural networks, has recently been used for biological image segmentation and classification (Moen *et al.*, 2019). The CellProfiler and DeepCell tools have recently been improved to include deep learning algorithms for the analysis of 3D images and complex patterns (McQuin *et al.*, 2018; Bannon *et al.*, 2018). However, their use in fluorescent pollen tetrad analysis is still limited, as software for tetrad analysis must integrate image information from single pollen grains and tetrads via a joint model with fluorescence patterns. Here, we have developed DeepTetrad (<https://github.com/abyslover/deeptetrad>), a deep learning-based image recognition package that enables quick, high-throughput, automated pollen tetrad analysis that can be used with existing FTL lines.

RESULTS

Establishment of DeepTetrad

To develop DeepTetrad, we adapted the Mask Regional Convolutional Neural Network (Mask R-CNN), integrating a deep residual network (ResNet) backbone for image recognition to detect four-pollen tetrads with and without fluorescence (He *et al.*, 2015; He *et al.*, 2017) (Figure 2). First, DeepTetrad must precisely recognize pollen tetrads in bright-field pollen images, which include not only tetrads but also triads, dyads and monads (Figure 2a–c). DeepTetrad was assembled with two separate Mask R-CNN processes, using a ResNet-feature pyramid network (FPN) backbone to generate masks of the bright-field pollen images (Figure 2a) (Lin *et al.*, 2016; He *et al.*, 2017). Then, DeepTetrad was trained to detect whole tetrad, triad, or dyad images via a Tetrad Segmentation Model with a ResNet depth of 101 layers. In parallel, we also trained DeepTetrad to detect every single pollen grain cell within tetrads, triads, dyads, and even monads via a Pollen Segmentation Model with a ResNet of depth 50 layers (Figure 2a). We used Keras and TensorFlow backends for training, with input bright-field images of pollen tetrads from FTLs (Chollet, 2015; Abadi *et al.*, 2016).

When trained, DeepTetrad can produce masks of both tetrad-like (tetrads, triads, dyads) and single pollen-like objects from bright-field images of pollen tetrads (Figure 2b). Next, DeepTetrad assigns a centroid to each pollen mask. Based on the position and distance between centroids of pollen masks within each tetrad-like mask, DeepTetrad recognizes measurable tetrads comprising four detectable pollen grains in the bright-field images (Figure 2c). DeepTetrad's tetrad classifier then determines a tetrad type from a choice of 12 classes (A–L) for three-color assays (Figures 1 and S3), or from a choice of three types (PD, T, NPD) for two-color assays (Figures S2 and S4),

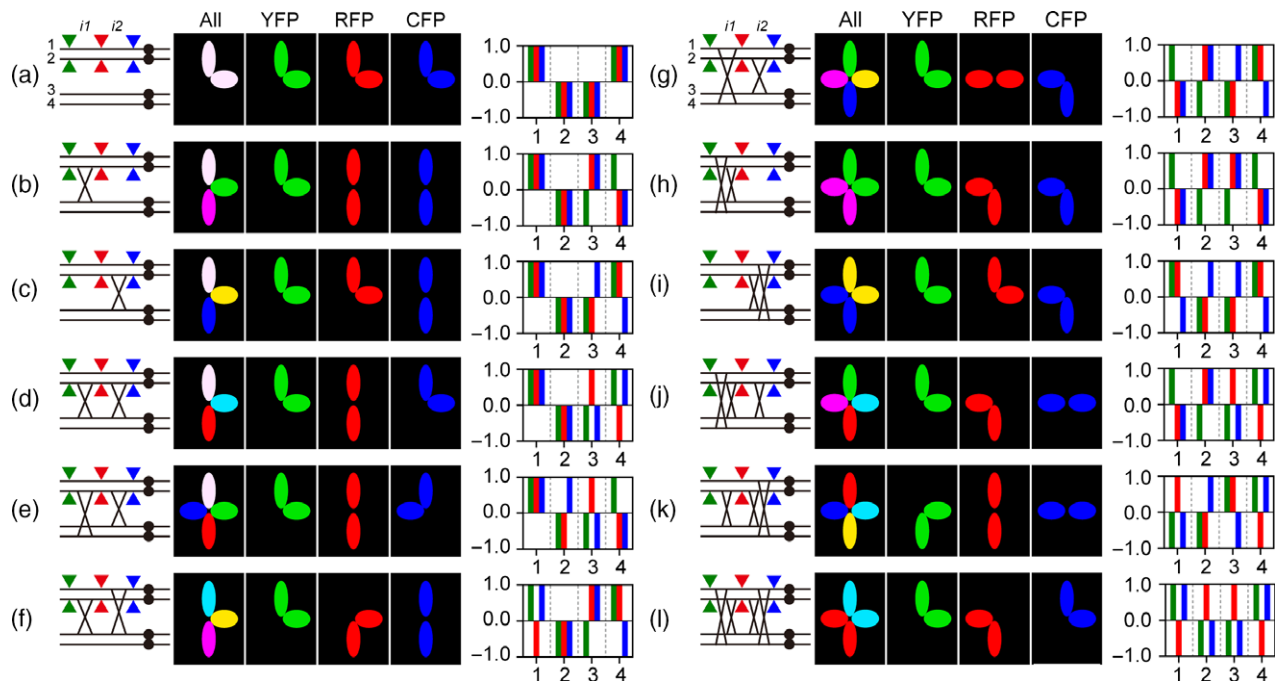


Figure 1. Diagram of 12 tetrad classes (A–L) generated by three-color tetrad assay and DeepTetrad outputs in the FTL system. Twelve tetrad classes (A–L) are generated from three-color tetrad assays (Berchowitz and Copenhaver, 2008). According to the position and segregation of three T-DNAs expressing eYFP, dsRed and eCFP, no recombination (A) and recombination tetrad classes (B–L) are determined in FTLs (*YRC+++*). The positions of COs between four chromatid strands (1st to 4st) in two intervals (*i1* and *i2*) are shown in B–L classes. Tetrad classes with CO are B, single crossover interval 1 SCO-*i1*, C, SCO-*i2*, D, two-strand double crossover 2st DCO, E, 3st DCOa, F, 3st DCOb, G, 4st DCO, H, NPD-*i1* NCO-*i2*, I, NCO-*i1* NPD-*i2*, J, NPD-*i1* SCO-*i2*, K, SCO-*i1* NPD-*i2*, and L, NPD-*i1* NPD-*i2*. For each tetrad class, all merged, green (YFP filter), red (RFP filter), and blue (CFP filter) images are shown. In the merged images, white indicates the superposition of green (YFP filter), red (RFP filter), and blue colors (CFP filter). Purple, yellow, and sky blue represent the superposition of red and blue, green and red, and green and blue, respectively. Bar graphs indicate the output of DeepTetrad classification. In the bar graphs, X-axis labels indicate four pollens (1, 2, 3, 4) per tetrad and Y-axis labels show the intensities of three-color fluorescence in the tetrad images.

according to the segregation pattern and intensity of fluorescence (yellow, red, cyan) in the four-pollen masks per tetrad mask (Figure 2d). CO frequency and interference can then be calculated using the frequency of tetrads in each class (Berchowitz and Copenhaver, 2008). Because DeepTetrad is able to recognize single pollen grains and classify their fluorescence in tetrads, triads, dyads and monads, we developed the DeepMonad package by subclassing DeepTetrad. Like flow cytometry analysis, DeepMonad can measure crossover frequency and interference in FTLs by analyzing images of single fluorescent pollen grains on the *qrt1/+* and *QRT1* backgrounds (Yelina *et al.*, 2013). In addition, DeepMonad can analyze single grains within tetrads, and this allows comparison of genetic distances and interference calculated by DeepTetrad and DeepMonad (Figure 3).

Validation and application of DeepTetrad

As DeepTetrad requires standard fluorescence microscopy, we developed a quick, simple method to prepare a large number of pollen tetrads for high-throughput imaging (Figure 3a). This method allows extensive image sets of pollen tetrads (bright-field, red, yellow and cyan) to be obtained, which can then be analyzed quickly and simultaneously by

DeepTetrad (Table 1). We used this technique to measure genetic distances in two-color FTL intervals (*CEN3*, *11b*, *11c*, *11b-c*, *11f*, *11g*, *11f-g*, *12f*, *12g*, *12f-g*, *13b*, *13c*, *13b-c*, *15a*, *15b*, *15a-b*) using DeepTetrad (Figures 3b and S1). The genetic distance values obtained this way were similar to those obtained using manual tetrad counting, flow cytometry and DeepMonad (Figure 3b and Tables 2, S1 and S2). Intriguingly, our DeepTetrad analysis showed higher crossover frequencies for long intervals (*11b*, *11c*, *11b-c*, *13b-c*, *15a*, *15b*, *15a-b*) compared with DeepMonad single pollen analysis; this was because DeepTetrad, but not DeepMonad, detects double crossovers in long intervals (Figure 3b and Tables 2, S1 and S2). In addition, the *CEN3* interval which spans the centromere on chromosome 3 had a lower CO rate (2.21 cM/Mb) than the overall male chromosome average CO frequencies (4.77 cM/Mb), and intervals *12f*, *12g* and *12f-g* which are close to the telomere had higher CO frequencies (10.19 cM/Mb, 10.71 cM/Mb and 10.23 cM/Mb, respectively) (Figures 3c and S1), consistent with prior observations (Giraut *et al.*, 2011; Choi *et al.*, 2013; Choi *et al.*, 2018). DeepTetrad can also recognize tetrad images taken at different magnifications and produces consistent CO frequency irrespective of scale (Figures 3d and S5).

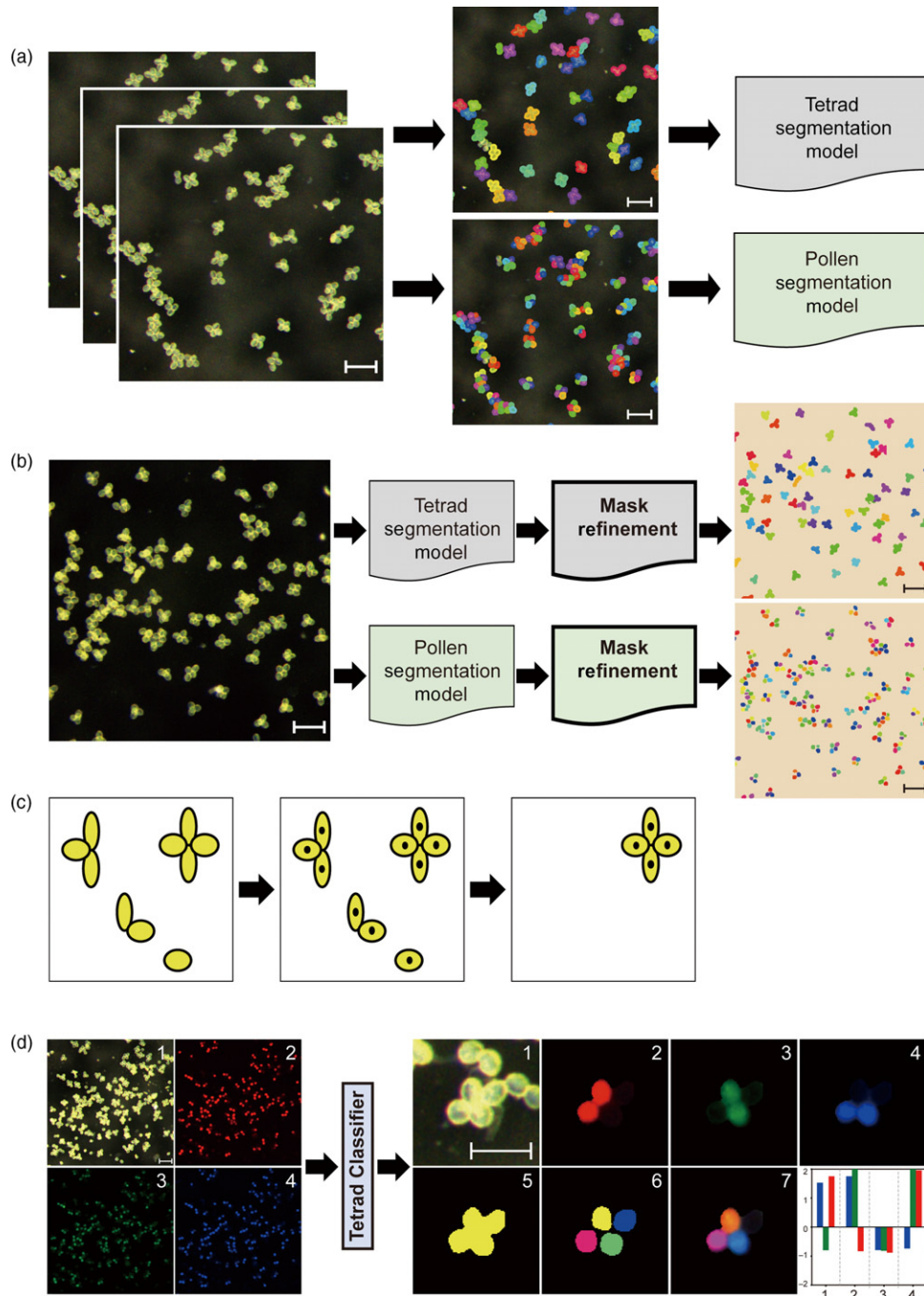


Figure 2. Establishment of DeepTetrad.

(a) Masking and training of tetrad-like and pollen images by DeepTetrad. Two separate DeepTetrad segmentation models make masks of tetrad-like and single pollens, respectively. (b) Generation of masks from tetrad-like and single-pollen images by DeepTetrad. (c) Recognition and selection of measurable tetrad masks by DeepTetrad. Black dots represent the centroid assigned to each pollen mask in monads, dyads, triads, and tetrads. (d) Tetrad classification by DeepTetrad. Bright-field (1), red (2), yellow (3), cyan (4) filtered tetrad images, tetrad mask (5), single-pollen masks (6), three-color merged tetrads (7) and DeepTetrad output are displayed. In the bar graphs of DeepTetrad output, X-axis labels indicate four pollens per tetrad and Y-axis labels show the intensities of three-color fluorescence in the tetrad image. Scale bar represents 0.1 mm (a, b, d, left), 0.05 mm (d, right). The colors in the tetrad mask images [a, middle panel, b, right panel, d, (6)] do not correspond to fluorescence colors.

To demonstrate DeepTetrad's utility for measuring CO interference we analyzed tetrad images from the three-color (*YRC/+++*) FTL interval-*11bc* (Figures 3e and S1).

Previously, interference had been measured in manually counted tetrads by calculating the interference ratio (defined here as σ) of the map distance of an interval (*i1*) in

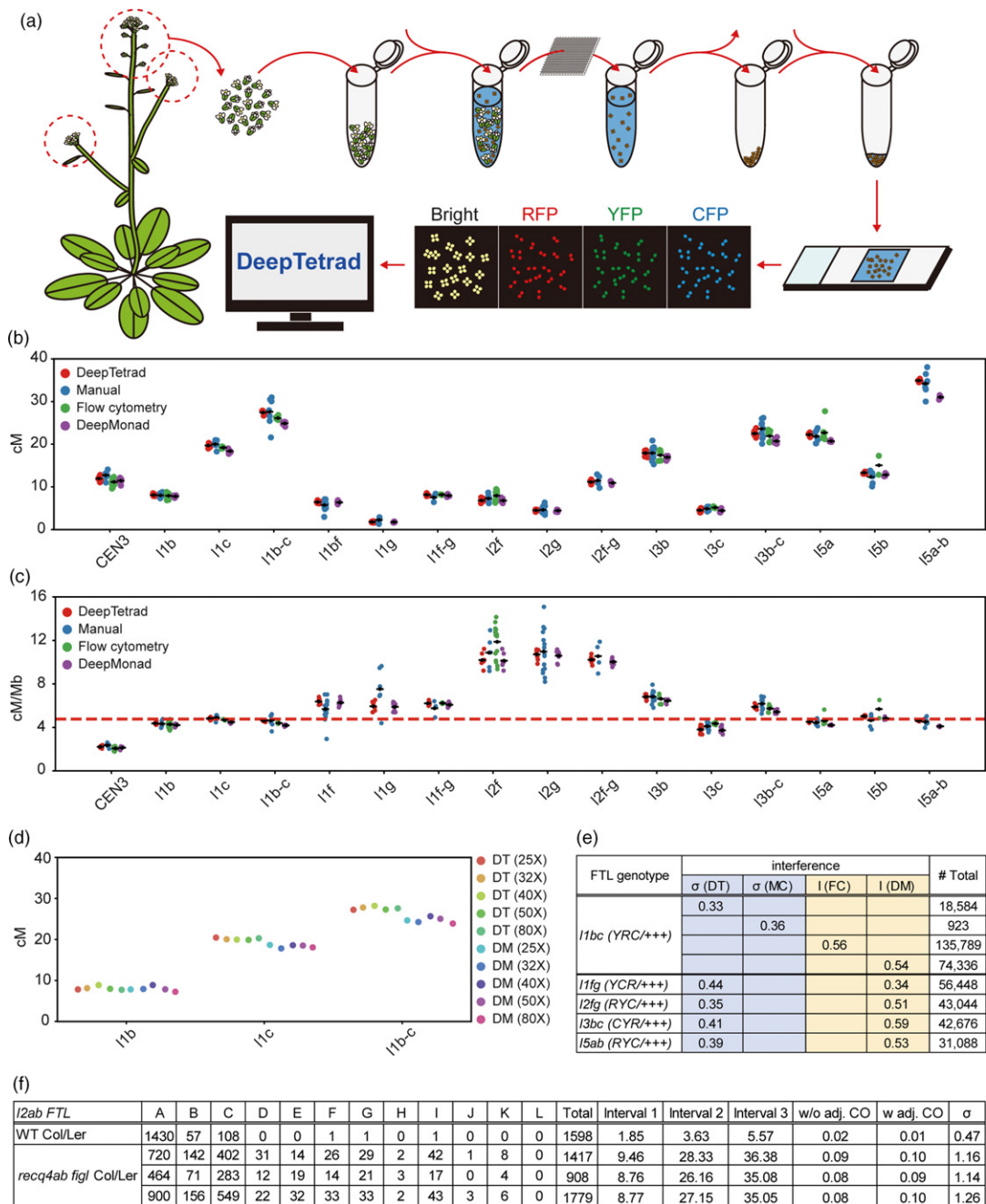


Figure 3. Measurements of crossover frequency and interference by DeepTetrad.

(a) A quick tetrad preparation method for high-throughput imaging of tetrads. The detail procedure is described in Experimental Procedures. (b) Plot showing measurement of CO frequencies (cM) in single intervals of FTLs. Genetic distances of single intervals were measured by DeepTetrad, manual counting, flow cytometry and DeepMonad. (c) Plot showing measurement of CO frequencies (cM/Mb) in single intervals of FTLs. A horizontal red line indicates the male chromosome average crossover rate. (d) Plot showing measurement of genetic distances in various sized tetrad images by DeepTetrad. Different magnifications were applied to the same tetrad samples for imaging (Figure S6). (e) Measurement of CO interference by DeepTetrad. The CO interference ratio ($\sigma = X_{i1}$ without adjacent CO/ X_{i1} with adjacent CO) was measured by DeepTetrad (DT) and manual counting (MC), highlighted in blue. Interference value ($I = 1 - \text{coefficient of coincidence}$) in yellow, was calculated by flow cytometry (FC) and DeepMonad (DM). A value of 1 and 0 indicates no interference in σ and I , respectively. The values of interference in other FTLs were measured by DeepTetrad and DeepMonad. (f) The CO interference ratio in *recq4a recq4b fig1* plants. DeepTetrad shows that *recq4a recq4b fig1* causes interference to be absent, increasing crossover frequency in FTL-*i2ab*.

tetrads that have a CO in an adjacent interval (*i2*) with the map distance of the same interval (*i1*) in tetrads that lack a CO in *i2* (Berchowitz and Copenhagen, 2008). Analysis of

18,584 tetrads with DeepTetrad resulted in a σ value of 0.33 for *11bc* which is consistent with the σ value of 0.36 obtained by manually counting 923 tetrads (Figure 3e and

Table 1 Comparison of crossover measurement methods

Comparison subject	Tetrad analysis (manual counting)	Flow cytometry	Tetrad analysis (DeepTetrad)
Equipment requirements	Fluorescence microscope, graphics software	Fluorescence microscope, flow cytometer	Fluorescence microscope, DeepTetrad package
Tetrad or pollen preparation method	Simple	Multiple steps (50-ml tube, flow cytometric tube)	Quick and simple (1.5-ml tube)
Time to prepare tetrads or pollens from 10 individual plants	2 h 30 min (1 h 40 min/ sampling, 50 min/imaging)	1 h (60 min/sampling)	1 h (10 min/sampling, 50 min/imaging)
Time of data analysis for two adjacent intervals in 10 individual plants (~1000 tetrads/plant)	30 h	5 h (30 min/plant, 10 plants)	2 h 30 min (16 min/plant, 10 plants)
Imaging requirement	Yes	No	Yes
Data analysis	Individually	Individually	Simultaneously
Gene conversion measurement	Yes	No	Yes
CO interference measurement	Yes	Yes	Yes
Single-interval DCO measurement	Yes	No	Yes
Differentiating 2-strand, 3-strand and 4-strand DCOs from one another	Yes	No	Yes
<i>qrt1</i> mutant background	Yes	No	Yes

DeepTetrad involves the preparation of a large number of tetrads using a quick and simple method, and increases the speed of tetrad analysis to obtain data from many individual plants by analyzing all tetrad images simultaneously. CO, crossover; DCO, double crossovers.

Tables 2, and S1–S3). Flow cytometry has also been used to measure interference in fluorescent-tagged pollen monads by calculating the ratio of observed double COs to expected double COs ($I = 1 - \text{coefficient of coincidence}$, the ratio of DCO_{obs} to DCO_{exp}) (Yelina *et al.*, 2013). An analysis of 74 336 monads (converted from tetrad images) using this method with DeepTetrad resulted in an interference value of 0.54 for *l1bc* (*YRC/+++*) which was consistent with a value of 0.56 obtained from flow sorting 135,789 monads (Figure 2e) (Yelina *et al.*, 2013). To provide baseline values for future studies we also used DeepTetrad and DeepMonad to measure σ and I values in four other three-color FTL intervals (*l1fg*, *l2fg*, *l3bc* and *l5ab*; (Figure 3e and Tables 2 and S1). Previously, it was shown that frequency of type II COs increases in *fancm* single mutants, as well as *recq4a recq4b fig1* triple mutants, leading to an absence of interference ($I = 0$, $\sigma = 1$) (Crismani *et al.*, 2012; Yelina *et al.*, 2013; Girard *et al.*, 2015; Fernandes *et al.*, 2018). Using DeepTetrad, we found that the σ value of FTL-*l2ab* in *recq4a recq4b fig1* plants is 1, indicating no detectable interference, consistent with the prior observations (Figure 3f and Tables 2 and S1).

Accuracy and sensitivity of DeepTetrad

Next, we examined how sensitively and specifically DeepTetrad recognizes single pollen grains and measurable tetrads. We compared bright-field images of single pollen grains ($n = 6975$, ~581/image) and tetrads ($n = 1721$, ~143/image) with their recognized masks by DeepTetrad in the *l1bc* interval (Tables S4–S8 and Figures S6 and S7). DeepTetrad detected single pollen grains ($n = 5452$, ~454/image)

and measurable tetrads ($n = 1363$, ~113/image) with a recognition ratio of approximately 0.797 and 0.806, respectively (Tables S4–S8). Through statistical binary classification analysis of the performance of DeepTetrad, we determined that DeepTetrad effectively predicts single pollen (sensitivity = 0.908 ± 0.022 , specificity = 0.986 ± 0.010 , precision = 0.985 ± 0.011 , accuracy = 0.947 ± 0.013) and tetrad masks (sensitivity = 0.950 ± 0.018 , specificity = 0.983 ± 0.013 , precision = 0.983 ± 0.013 , accuracy = 0.966 ± 0.012) by jointly mapping two separate masks in a single plane (Table S4). We also determined that each FTL has a distinct level of silenced fluorescent pollen grains on the *qrt1/+* background using CellProfiler (Table S9). Manual tetrad counting and flow cytometric FTL analysis are affected by the presence of silenced fluorescent pollen. Tetrads with silenced fluorescent pollen grains must be manually removed in the original FTL analysis protocol, and the ratio of fluorescent to non-fluorescent pollen populations must be considered when performing in flow cytometric FTL analysis (Berchowitz and Copenhaver, 2008; Yelina *et al.*, 2013). However, DeepTetrad can recognize and classify images of tetrads with two fluorescent and two non-fluorescent pollen grains in each fluorescence channel (Tables S7 and S8). Taken together, our data demonstrate that DeepTetrad is a useful deep learning-based image recognition package for high-throughput measurements of both CO frequency and interference.

DISCUSSION

The FTL-based visual tetrad assay has been used extensively in studies of plant meiosis (Francis *et al.*, 2007;

Table 2 Measurement of crossover frequency of FTL intervals by DeepTetrad, manual counting, flow cytometry, and DeepMonad

FTL genotype	T-DNA 1	T-DNA 2	T-DNA 3	Mb	cM (DT)	cM (MC)	cM (FC)	cM (DM)	P-value (DT-MC)	P-value (DT-DM)	Total tetrads (DT)
<i>l1bc (YRC/+++)</i>	3 905 441	5 755 618	9 850 022								18 584
<i>l1b (YR/++)</i>	3 905 441	5 755 618		1.85	8.09 ± 0.17	8.02 ± 0.59	7.93 ± 0.18	7.79 ± 0.20	7.68E-01	1.54E-02*	
<i>l1c (RC/++)</i>		5 755 618	9 850 022	4.09	19.69 ± 0.37	20.00 ± 0.83	19.20 ± 0.21	18.38 ± 0.32	3.97E-01	2.02E-05*	
<i>l1b-c (YC/++)</i>	3 905 441		9 850 022	5.94	27.43 ± 0.36	27.59 ± 3.10	26.12 ± 0.42	24.88 ± 0.32	9.07E-01	5.29E-09*	
<i>l1fg (YCR/+++)</i>	24 645 163	25 652 977	25 956 590								14 112
<i>l1f (YC/++)</i>	24 645 163	25 652 977		1.01	6.44 ± 0.21	5.74 ± 0.5		6.34 ± 0.26	5.61E-02	4.78E-01	
<i>l1g (CR/++)</i>		25 652 977	25 956 590	0.30	1.78 ± 0.11	2.26 ± 0.61		1.77 ± 0.10	1.02E-01	8.19E-01	
<i>l1f-g (YR/++)</i>	24 645 163		25 956 590	1.31	8.14 ± 0.18	7.54 ± 2.5	8.16 ± 0.07	7.95 ± 0.18	1.16E-01	1.05E-01	
<i>l2fg (RYC/+++)</i>	18 286 716	18 957 093	19 373 634								10 761
<i>l2f (RY/++)</i>	18 286 716	18 957 093		0.67	6.83 ± 0.33	7.29 ± 1.3	7.95 ± 0.58	6.79 ± 0.34	3.94E-01	8.63E-01	
<i>l2g (YC/++)</i>		18 957 093	19 373 634	0.42	4.50 ± 0.17	4.61 ± 0.46		4.45 ± 0.18	6.31E-01	6.08E-01	
<i>l2f-g (RC/++)</i>	18 286 716		19 373 634	1.09	11.15 ± 0.28	11.5 ± 2.31		10.93 ± 0.26	6.67E-01	1.93E-01	
<i>l3bc (CYR/+++)</i>	498 916	3 126 994	4 319 513								10 669
<i>l3b (CY/++)</i>	498 916	3 126 994		2.63	17.91 ± 0.54	17.94 ± 0.71	17.43 ± 1.14	16.95 ± 0.35	9.58E-01	NA	
<i>l3c (YR/++)</i>		3 126 994	4 319 513	1.19	4.53 ± 0.35	4.88 ± 0.28	5.17 ± 0.14	4.44 ± 0.31	8.17E-02	6.58E-01	
<i>l3b-c (CR/++)</i>	498 916		4 319 513	3.82	22.51 ± 0.64	23.61 ± 1.46	21.89 ± 1.22	20.72 ± 0.53	1.44E-01	1.65E-04*	
<i>l5ab (RYC/+++)</i>	18 164 269	23 080 567	25 731 311								7772
<i>l5a (RY/++)</i>	18 164 269	23 080 567		4.92	22.24 ± 0.49	21.86 ± 1.36	22.70 ± 1.77	20.73 ± 0.27	5.32E-01	7.28E-05*	
<i>l5b (YC/++)</i>		23 080 567	25 731 311	2.65	13.31 ± 0.31	12.36 ± 1.28		12.81 ± 0.42	1.62E-01	2.68E-02*	
<i>l5a-b (RC/++)</i>	18 164 269		25 731 311	7.57	34.91 ± 0.38	34.21 ± 2.41		31.03 ± 0.50	5.05E-01	1.35E-07*	
<i>CEN3 (RY/+++)</i>	11 115 724	16 520 560		5.40	11.96 ± 0.64	12.72 ± 0.86	11.17 ± 0.19	11.45 ± 0.70	1.16E-01	1.95E-01	7657
<i>l2ab recq4ab figl (CYR/+++)</i>	12 640 092	13 226 013	14 675 407								4271
<i>l2a recq4ab figl (CY/++)</i>	12 640 092	13 226 013		0.59	8.86 ± 1.38			8.35 ± 1.56	NA	NA	
<i>l2b recq4ab figl (YR/++)</i>		13 226 013	14 675 407	1.45	27.47 ± 1.88			21.42 ± 0.48	NA	NA	
<i>l2a-b recq4ab figl (CR/++)</i>	12 640 092		14 675 407	2.04	35.26 ± 2.56			25.99 ± 0.83	NA	NA	

In FTLs with T-DNAs expressing eYFP (Y), dsRed (R) and eCFP (C), the physical distances and genetic distances are shown (Berchowitz and Copenhaver, 2008). Genetic distances were measured by DeepTetrad, manual counting, flow cytometry (Yelina *et al.*, 2013; Ziolkowski *et al.*, 2015; Yelina *et al.*, 2015; Ziolkowski *et al.*, 2017; Choi *et al.*, 2018), and DeepMonad. The results of statistical analyses (mean of cM, 95% confidence interval, P-value) on the genetic distances measured by different methods are shown. The P-values of significantly different crossover frequency are marked by asterisks (*). A P-value is not calculated when assumptions for the statistical t-test are violated. DT, DeepTetrad; MC, manual counting; FC, flow cytometry; DM, DeepMonad; FTL, fluorescent-tagged line.

Berchowitz and Copenhaver, 2008; Fernandes *et al.*, 2018). The application of flow cytometry to FTLs allowed rapid, high-throughput measurement of CO frequency and interference (Yelina *et al.*, 2013; Ziolkowski *et al.*, 2017). Both the original FTL visual assay and flow cytometric FTL analysis have clear benefits and weaknesses. The original FTL assay requires manual counting using fluorescence microscopy and is relatively slow and labor intensive. Flow cytometric FTL analysis is more rapid and reproducible but requires a multi-color flow cytometer and high-purity pollen grains. Here, we have extended the utility of FTLs further by developing DeepTetrad to enable quick, simple, automated tetrad analysis. The DeepTetrad system builds upon the benefits of both visual and flow cytometric FTL analysis and overcomes their weaknesses by improving the ease of use, speed, and accuracy of the FTL system.

The DeepTetrad system has the following advantages and potential applications to facilitate a wide array of studies: (i) DeepTetrad saves time and labor using high-throughput tetrad preparation, imaging, and automatic computational analysis (Table 1); (ii) DeepTetrad has enhanced reproducibility and accuracy by automatically excluding the individual variability inherent in manual tetrad counting and the effects of fluorescence silencing; (iii) DeepTetrad does not require the use of a sophisticated flow cytometer or high-purity pollen grains, instead it uses a rapid tetrad preparation method; (iv) DeepTetrad enables the detection of double COs by keeping the FTL in the *qrt* background; (v) In the future we plan to modify DeepTetrad to measure gene conversion rates using FTLs by changing the DeepTetrad classifier from tetrad classification mode to the detection of gene conversions (Francis *et al.*, 2007); (vi) DeepTetrad also has the potential to be used for forward genetic screens of altered CO frequency and interference mutants; (vii) DeepTetrad measures CO frequency in defined intervals along the genome, whereas whole-genome sequencing can generate a genome-wide CO map at high resolution. However, DeepTetrad can help researchers quickly decide whether to generate genome-wide CO maps in diverse genetic backgrounds, a process that is expensive and time consuming; (viii) We anticipate that it will be possible to adapt it for use in other fluorescence-based tetrad analysis systems, including those that have been developed for *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (Thacker *et al.*, 2011; Li *et al.*, 2019). Importantly, the high speed and accuracy of DeepTetrad result from its ability to recognize and classify tetrad images from mixtures of monads, dyads, triads, and tetrads. Hence, DeepTetrad will accelerate genetic analysis of meiotic recombination mechanisms, as well as the influence of epigenetic and environmental effects.

EXPERIMENTAL PROCEDURES

DeepTetrad network architecture

DeepTetrad assembles two separate Mask Regional Convolutional Neural Network (Mask R-CNN) for the instance segmentation task (He *et al.*, 2017). They generate masks of pollen objects and tetrad objects, respectively, from input bright-field images. As backbone architectures, which are responsible for feature detection, we used deep residual networks (ResNet) of depths 50 and 101, with a FPN (Girshick, 2015). We use the same terminology and definitions as those used in the Mask R-CNN article (He *et al.*, 2017) when describing the backbone; ResNet-50-FPN. Multi-task loss L is also defined in the same manner:

$$L = L_{\text{cls}} + L_{\text{box}} + L_{\text{mask}},$$

where L_{cls} is classification loss. L_{box} is loss of bounding box $v = (v_x, v_y, v_w, v_h)$, which is a rectangle defined by coordinates of the upper-top vertex (x, y) , and the dimension (width w , height h). L_{mask} is mask loss. The i -th mask of bounding box v , which is the core feature in DeepTetrad, is a grayscale image defined by the following logical predicate:

$$m_i(v) = \{(x_i, y_i) | \text{pixel}(x_i, y_i) > 0, v_x \leq x_i < v_x + v_w, v_y \leq y_i < v_y + v_h, i > 0\},$$

in which the $\text{pixel}(x, y)$ function returns the pixel value of given coordinates in an image.

Mask segmentation is a multi-label classification. Hence L_{mask} must be calculated independently for each class in a single image. It is achieved by applying a sigmoid function to each pixel, from which the mean of binary cross entropy loss is calculated:

$$L_{\text{mask}} = - \sum_{i=1}^{C=2} y_i \log \left(\frac{1}{1 + e^{x_i}} \right).$$

Training

DeepTetrad is trained by nVidia TITAN X with 12 GB RAM using Keras (Chollet, 2015) and Tensorflow (Abadi *et al.*, 2016) backends in the CUDA 10 platform. Transfer learning is performed with pre-trained weights on a Microsoft COCO dataset (Lin *et al.*, 2015). Input images are of fixed dimensions (1920, 2560). Zero-padding resizes each image to the exact dimension of (2048, 2560), which ensures that width and height are multiples of 512. The image is cropped at random positions with dimensions of (512, 512). For pollen objects, 919 training masks and 370 validation masks were used for data augmentation. For tetrad objects, 1371 training masks and 617 validation masks were used. Masks were annotated using VGG Image Annotator (VIA) (Dutta and Zisserman, 2019).

Image augmentation with scaling, translation, rotation, and shearing operations is randomly triggered to each training session and validation mini-batch. During an epoch, a mini-batch of two images per Graphic Processing Unit (GPU) is fed to the backbone. Regions of interest (ROIs) or bounding boxes are sampled 128 times for the pollen model; 512 times for the tetrad model.

The pollen model, in which ResNet-50-FPN backbone is integrated, is trained by a single GPU for 10 000 iterations using the Stochastic Gradient Descent optimizer with a learning rate of 0.001, momentum of 0.9, and weight decay of 0.0001. The tetrad model is trained with the same configuration, except that the

backbone is replaced with ResNet-101-FPN and the number of iterations is increased to 20 000.

Inference

Each backbone applies non-maximum suppression to 6000 ROI candidates, in turn yielding 1000 ROIs. The number of detected masks is limited by the maximum number of detecting instances D , which is set to 200. Hence, only D -detected ROIs with the highest scores are selected to create masks. If D is increased, the model may fail to infer masks because of memory limitation, or the inference may be seriously prolonged. The model might also overlook a considerable number of masks, which would be a critical problem.

As a solution, DeepTetrad tries to infer masks from cropped images rather than directly gathering them from a whole image. In an image of dimension (2048, 2560), there are at least 1600–5000 ground-truth pollen masks, and at least 400–1250 ground-truth tetrad masks. As well as tetrad masks, monad, dyad, and triad masks will also be reported by the tetrad model. Thus, the number of ground-truth masks in the whole image is much larger than D in both cases. In total, 63 cropped images of dimension (512, 512) are generated from the whole image, thereby remaining images of dimension (256, 512), or top images of dimension (512, 256) intersect with one another. Complete inference for the whole image involves predicting masks from a mini-batch of 21 cropped images in three epochs.

Mask refinement

All masks from the inference stage need to be refined. Masks that are produced at the edges of the cropped image are usually not overlapping in local coordinates, which represent the spatial location before translating to that in the whole image. However, they can become broken or overlapping when translated to global coordinates, which are the coordinates in the whole image. After removing broken or overlapping masks, contours and the number of tetrad and pollen masks can be more accurately determined. The procedure is as follows: (i) all pixel coordinates of pollen masks, M_p , and tetrad masks, M_t in each cropped image are moved to global coordinates by translation (affine transformation). (ii) All pixel coordinates of tetrad masks are stored in a k -d tree, T_t . Similarly, all pollen masks are kept in a k -d tree, T_p . (iii) All tetrad masks are queried against T_t and, similarly, all pollen masks are queried against T_p to collect refined tetrad masks, Ψ_t , and refined pollen masks, Ψ_p , which meet the predicate below:

$$\Psi_c(M_c) = \{m_i | \text{dist}(m_i, m_j) = 0, n(m_i) > n(m_j), i \neq j, m_i \in M_c, m_j \in M_c\},$$

in which, $\text{dist}(x, y)$ returns a Euclidean distance between x and y , $n(x)$ returns the number of elements, and c is either t (tetrad) or p (pollen).

The measurable tetrad masks and pollen masks, Ω , are defined as follows:

$$\Omega(\Psi_t, \Psi_p) = \{(\psi_i, \psi_j) | n(\psi_i \cap \text{centroid}(\psi_j)) = 4, \psi_i \in \Psi_t, \psi_j \in \Psi_p\},$$

where $\text{centroid}(x)$ yields the median of given mask coordinates.

Centroids are calculated to associate a tetrad mask Ψ_t with the refined pollen masks Ψ_p . Each centroid of a pollen mask, Ψ_p , is queried against T_t , then Ψ_p is associated with a tetrad mask, Ψ_t if the Euclidean distance between them is 0. When the number of Ψ_p associated with Ψ_t is four, they are deemed to be measurable.

Tetrad classification

A tetrad mask can be classified into a representative type of crossover event according to the fluorescence intensity values of associated pollen masks. The signal intensity, S_c , is defined as the mean of pixel values in each fluorescence channel of measurable pollen masks, in which c can be a fluorescence channel of red (R), yellow (Y), or cyan (C). We assume $c = \{R, Y\}$ for two-channel images, and $c = \{R, Y, C\}$ for three-channel images. For measurable tetrad masks, S_c of four-pollen masks can be calculated, yet those which have undergone silencing in fluorescent protein expression should be ignored. If silencing occurs, the difference between S_c of the second-highest and the third-highest would be smaller than a certain threshold, Θ . Second-highest S_c represents presence or on-state of fluorescence proteins, whereas the third-highest implies absence or off-state of the proteins. Z-scores are calculated for four S_c values before finding Θ . We explored a parameter space of Θ up to two decimal places using an adaptive grid search method, then Θ was set to 0.40, meaning that S_c differences between on-state and off-state must be bigger than but not equal to 0.40 in a normal distribution.

We determine if fluorescent proteins in a pollen grain are expressed by comparing individual S_c with the median of all four S_c values. With per-channel expressions, tetrad masks are classified as one of classical three tetrad types: PD, tetra type (T), and NPD, with two-color FTL intervals or 12 classes (A to L) with the three-color counterparts (Figures 1 and S2). In three-color FTL intervals that have two intervals ($i1$ and $i2$) with four chromatids (1–4), tetrad classes are no recombination (A), single crossover interval 1 (B; SCO- $i1$), single crossover interval 2 (C; SCO- $i2$), two-strand double crossover (D; 2stDCO), three-strand double crossover a (E; 3st DCOa), three-strand double crossover b (F; 3st DCOb), four-strand double crossover (G; 4st DCO), non-parental ditype interval 1, non-crossover interval 2 (H; NPD- $i1$ NCO- $i2$), non-crossover interval 1, non-parental ditype interval 2 (I; NCO- $i1$ NPD- $i2$), non-parental ditype interval 1, single crossover interval 2 (J; NPD- $i1$ SCO- $i2$), single crossover interval 1, non-parental ditype interval 2 (K; SCO- $i1$ NPD- $i2$) and non-parental ditype interval 1, non-parental ditype interval 2 (L; NPD- $i1$ NPD- $i2$) (Berchowitz and Copenhaver, 2008).

Calculation of interference

With two-color FTL intervals, we calculate crossover frequency following Perkin's equation:

$$cM = \frac{0.5T + 3NPD}{(PD + T + NPD)} * 100.$$

With three-color FTL intervals, we can calculate the interference ratio σ , which is the ratio of the map distance with adjacent crossover χ_γ to the map distance without adjacent crossover χ_δ :

$$\chi_\gamma = \frac{0.5T_\gamma + 3NPD_\gamma}{PD_\gamma + T_\gamma + NPD_\gamma} = \frac{0.5(D + E + F + G + K) + 3(J + L)}{(C + I) + (D + E + F + G + K) + (J + L)}$$

$$\chi_\delta = \frac{0.5T_\delta + 3NPD_\delta}{PD_\delta + T_\delta + NPD_\delta} = \frac{0.5(B) + 3(H)}{(A) + (B) + (H)}$$

$$\sigma = \frac{\chi_\gamma}{\chi_\delta},$$

in which PD means the number of parental ditypes or no crossover event. T is the number of tetra type or single crossover events. NPD is the number of non-parental ditype or double

crossover events. A–L letters represent 12 tetrad classes from three-color FTLs (Figure 1) (Berchowitz and Copenhaver, 2008).

In the above equations, we assumed that two adjacent genetic intervals of $i1$ and $i2$ are defined by three separate fluorescent protein transgenes of red, yellow, and cyan in sequential order. The γ represents that at least a single crossover event occurred at $i2$, meanwhile δ denotes that no crossover events were found at $i2$. We highlight that the number of crossover events is counted at $i1$. T_γ is the tetratype tetrads for $i1$ that have a CO in $i2$, and T_δ is the tetratype tetrads for $i1$ that do not have a CO in $i2$. DeepTetrad maps the order of input color images (red–yellow–cyan) to the physical order of fluorescent protein transgenes of FTLs for calculating the interference ratio as well as genetic distance.

Pollen tetrad preparation

FTL plants were grown at 20°C under long-day conditions (16 h light/8 h dark). Twenty open flowers of a primary shoot from 30-day-old FTL plants were collected in a 1.5-ml tube, and 1 ml of pollen tetrad preparation solution (17% sucrose, 2 mM CaCl₂, 1.625 mM boric acid, 0.1% Triton-X-100, pH 7.5) was added before incubating for 5 min at room temperature, with gentle rotation. Flowers and the solution were mixed by inverting the tube several times. The solution of pollen tetrads was pipetted and filtered into a new 1.5-ml tube through an 80- μ m nylon mesh (30 x 30 mm). The filtered solution was centrifuged at 500 *g* for 3 min to make a yellow pellet. The supernatant was removed and discarded by pipetting or vacuum aspiration. Four μ l of pollen tetrad preparation solution was added to the yellow pellet. After pipetting gently five times, the 4- μ l suspension of pollen tetrads was loaded on a glass microscope slide and covered with a small cover glass (9 x 9 mm). This resulted in ~2500 tetrads for imaging.

Microscopy and imaging

A set of four photographs for each pollen tetrad was taken using a Leica M165 flow cytometry (FC) dissecting stereomicroscope with bright-field, RFP, YFP and CFP filters (Article Nos. 10450224, 10447410, 10447409, respectively) in sequential order. Twelve image sets per cover glass were obtained from ~20 flowers when a magnification of $\times 50$ was used to image tetrads. Information about the gain, gamma, saturation and exposure for each FTL for high quality imaging is available (Table S10).

ACKNOWLEDGMENTS

We thank Raphael Mercier (MPI, Germany) for providing *recq4 figl1* mutant seeds. We thank our colleagues for giving critical comments. This work is funded by the Suh Kyungbae Foundation, Next-Generation BioGreen 21 Program (PJ01337001), Rural Development Administration, Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1D1A03028374), and the Cooperative Research Program for Agriculture Science and Technology Development (PJ010953022019), Republic of Korea. GPC is supported by a US National Science Foundation grant (IOS-1844264).

AUTHOR CONTRIBUTIONS

E-CL, JaK, JuK, and KC contributed to design of the study. E-CL generated the computational code of DeepTetrad. JaK, and JP performed experiments. E-CL, JaK, JP, E-J K, JuK, YMP, HSC, DB, IRH, GPC, IH, and KC analyzed the

data. E-CL, JaK, JP, E-J K, JuK, IRH, GPC, IH, and KC wrote and revised the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DATA STATEMENT

All information of the DeepTetrad software package used for image and data analysis is available at Github (<https://github.com/abyslover/deeptetrad>).

SUPPORTING INFORMATION

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

Figure S1. T-DNA locations of pollen FTLs (*I1bc*, *I1fg*, *I2ab*, *I2fg*, *I3bc*, *CEN3*, *I5ab*) on the *Arabidopsis thaliana* genome.

Figure S2. Diagram for three tetrad classes (PD, T, NPD) of 2-color assay and DeepTetrad output.

Figure S3. Tetrad images and DeepTetrad output in three-color assay of FTL plants.

Figure S4. Tetrad images and DeepTetrad output in two-color assay of FTL-*CEN3* (*YR/++*) plants.

Figure S5. Different sized tetrad images.

Figure S6. Receiver operating characteristic (ROC) curve of tetrad classification.

Figure S7. Receiver operating characteristic curve of pollen classification.

Table S1. Measurements of crossover frequency and interference in FTL intervals by DeepTetrad.

Table S2. Measurements of crossover frequency in FTL intervals by manually counting tetrads.

Table S3. Measurements of crossover interference in FTL-*I1bc* by manually counting tetrads.

Table S4. Statistical analysis of the performance of the DeepTetrad models.

Table S5. Analysis of the performance of the Tetrad Segmentation Model.

Table S6. Analysis of the performance of the Pollen Segmentation Model.

Table S7. Analysis of silencing effects in the Tetrad Segmentation Model.

Table S8. Analysis of silencing effects in the Pollen Segmentation Model.

Table S9. Ratio of fluorescence-silenced pollen grains in FTLs.

Table S10. Information of imaging fluorescent pollen in FTLs.

REFERENCES

- Abadi, M., Barham, P., Chen, J. *et al.* (2016) TensorFlow: a system for large-scale machine learning. In *Operating Systems Design and Implementation*, 265–283. <https://arxiv.org/abs/1603.04467>
- Bannon, D., Moen, E., Borba, E., Ho, A., Camplisson, I., Chang, B., Osterman, E., Graf, W. and Van Valen, D. (2018) DeepCell 2.0: Automated cloud deployment of deep learning models for large-scale cellular image analysis. *bioRxiv*, <https://doi.org/10.1101/505032>.
- Berchowitz, L.E. and Copenhaver, G.P. (2008) Fluorescent Arabidopsis tetrads: a visual assay for quickly developing large crossover and crossover interference data sets. *Nat. Protoc.* **3**, 41–50.

- Berchowitz, L.E., Francis, K.E., Bey, A.L. and Copenhaver, G.P. (2007) The role of AtMUS81 in interference-insensitive crossovers in *A. thaliana*. *PLoS Genet.* **3**, e132.
- Carpenter, A.E., Jones, T.R., Lamprecht, M.R. *et al.* (2006) Cell Profiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100.
- Choi, K. (2017) Advances towards controlling meiotic recombination for plant breeding. *Mol. Cells*, **40**, 814–822.
- Choi, K. and Henderson, I.R. (2015) Meiotic recombination hotspots - a comparative view. *Plant J.* **83**, 52–61.
- Choi, K., Zhao, X., Kelly, K.A. *et al.* (2013) Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nat. Genet.* **45**, 1327–1336.
- Choi, K., Zhao, X., Tock, A.J. *et al.* (2018) Nucleosomes and DNA methylation shape meiotic DSB frequency in *Arabidopsis thaliana* transposons and gene regulatory regions. *Genome Res.* **28**, 532–546.
- Chollet, F. (2015) keras. *GitHub*. Available at: <https://github.com/fchollet/keras>
- Crismani, W., Girard, C., Froger, N., Pradillo, M., Santos, J.L., Chelysheva, L., Copenhaver, G.P., Horlow, C. and Mercier, R. (2012) FANCM limits meiotic crossovers. *Science*, **336**, 1588–1590.
- Dutta, A. and Zisserman, A. (2019) The VGG Image Annotator (VIA). *ArXiv*, <https://arxiv.org/abs/1904.10699>.
- Fernandes, J.B., Séguéla-Arnaud, M., Larchevêque, C., Lloyd, A.H. and Mercier, R. (2018) Unleashing meiotic crossovers in hybrid plants. *Proc. Natl. Acad. Sci. USA*, **115**, 2431–2436.
- Francis, K.E., Lam, S.Y., Harrison, B.D., Bey, A.L., Berchowitz, L.E. and Copenhaver, G.P. (2007) Pollen tetrad-based visual assay for meiotic recombination in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, **104**, 3913–3918.
- Girard, C., Chelysheva, L., Choinard, S., Froger, N., Macaisne, N., Lehmemdi, A., Mazel, J., Crismani, W. and Mercier, R. (2015) AAA-ATPase FIDGETIN-LIKE 1 and helicase FANCM antagonize meiotic crossovers by distinct mechanisms. *PLoS Genet.* **11**, e1005369.
- Giraut, L., Falque, M., Drouaud, J., Pereira, L., Martin, O.C. and Mézard, C. (2011) Genome-wide crossover distribution in Arabidopsis thaliana meiosis reveals sex-specific patterns along chromosomes. *PLoS Genet.* **7**, e1002354.
- Girshick, R. (2015) Fast R-CNN. *ArXiv*, **1504**, 08083.
- He, K., Zhang, X., Ren, S. and Sun, J. (2015) Deep residual learning for image recognition. *ArXiv*, **1512**, 03385.
- He, K., Gkioxari, G., Dollár, P. and Girshick, R. (2017) Mask R-CNN. *ArXiv*, **1703**, 06870.
- Higgins, J.D., Armstrong, S.J., Franklin, F.C.H. 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.
- Hunter, N. (2015) Meiotic recombination: the essence of heredity. *Cold Spring Harb. Perspect. Biol.* **7**, a016618.
- Li, D., Roca, M., Yucel, R. and Lorenz, A. (2019) Immediate visualization of recombination events and chromosome segregation defects in fission yeast meiosis. *Chromosoma*. <https://doi.org/10.1007/s00412-019-00691-y>.
- Lin, T.-Y., Maire, M., Belongie, S. *et al.* (2015) Microsoft COCO: Common Objects in Context. *ArXiv*, <https://arxiv.org/abs/1405.0312>.
- Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2016) Feature pyramid networks for object detection. *ArXiv*, **1612**, 03144.
- McQuin, C., Goodman, A., Chernyshev, V. *et al.* (2018) Cell profiler 3.0: next-generation image processing for biology. T. Misteli, ed. *PLoS Biol.* **16**, e2005970.
- 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. USA*, **109**, E981–E988.
- Melamed-Bessudo, C., Yehuda, E., Stuitje, A.R. and Levy, A.A. (2005) A new seed-based assay for meiotic recombination in Arabidopsis thaliana. *Plant J.* **43**, 458–466.
- Melamed-Bessudo, C., Shilo, S. and Levy, A.A. (2016) Meiotic recombination and genome evolution in plants. *Curr. Opin. Plant Biol.* **30**, 82–87.
- Mercier, R., Jolivet, S., Vezon, D. *et al.* (2005) Two meiotic crossover classes cohabit in Arabidopsis: one is dependent on MER3, whereas the other one is not. *Curr. Biol.* **15**, 692–701.
- Mercier, R., Mézard, C., Jenczewski, E., Macaisne, N. and Grelon, M. (2015) The molecular biology of meiosis in plants. *Annu. Rev. Plant Biol.* **66**, 297–327.
- Mieulet, D., Aubert, G., Bres, C. *et al.* (2018) Unleashing meiotic crossovers in crops. *Nat. Plants*, **4**, 1010–1016.
- Moën, E., Bannon, D., Kudo, T., Graf, W., Covert, M. and Valen, D. (2019) Deep learning for cellular image analysis. *Nat. Methods*. <https://doi.org/10.1038/s41592-019-0403-1>.
- Perkins, D.D. (1949) Biochemical mutants in the Smut Fungus *Ustilago Maydis*. *Genetics*, **34**, 607–626.
- Perkins, D.D. (1962) Crossing-over and interference in a multiply marked chromosome arm of Neurospora. *Genetics*, **47**, 1253–1274.
- Serra, H., Lambing, C., Griffin, C.H. *et al.* (2018) Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis. *Proc. Natl. Acad. Sci. USA*, **115**, 2437–2442.
- Shilo, S., Melamed-Bessudo, C., Dorone, Y., Barkai, N. and Levy, A.A. (2015) DNA crossover motifs associated with epigenetic modifications delineate open chromatin regions in Arabidopsis. *Plant Cell*, **27**, 2427–2436.
- Thacker, D., Lam, I., Knop, M. and Keeney, S. (2011) Exploiting spore-autonomous fluorescent protein expression to quantify meiotic chromosome behaviors in *Saccharomyces cerevisiae*. *Genetics*, **189**, 423–439.
- van Tol, N., Rolloos, M., van Loon, P. and van der Zaal, B.J. (2018) MeioSeed: a Cell Profiler-based program to count fluorescent seeds for crossover frequency analysis in Arabidopsis thaliana. *Plant Methods*, **14**, 32.
- Twell, D., Yamaguchi, J. and McCormick, S. (1990) Pollen-specific gene expression in transgenic plants: coordinate regulation of two different tomato gene promoters during microsporogenesis. *Development*, **109**, 705–713.
- Underwood, C.J., Choi, K., Lambing, C. *et al.* (2018) Epigenetic activation of meiotic recombination near *Arabidopsis thaliana* centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome Res.* **28**, 519–531.
- Vrielynck, N., Chambon, A., Vezon, D., Pereira, L., Chelysheva, L., Muyt, A.De, Mézard, C., Mayer, C. and Grelon, M. (2016) A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science*, **351**, 939–943.
- Wang, Y. and Copenhaver, G.P. (2018) Meiotic recombination: mixing it up in plants. *Annu. Rev. Plant Biol.* **69**, 577–609.
- Wijnker, E., Velikkam James, G., Ding, J. *et al.* (2013) The genomic landscape of meiotic crossovers and gene conversions in Arabidopsis thaliana. *Elife*, **2**, e01426.
- Yelina, N.E., Choi, K., Chelysheva, L. *et al.* (2012) Epigenetic remodeling of meiotic crossover frequency in Arabidopsis thaliana DNA methyltransferase mutants. *PLoS Genet.* **8**, e1002844.
- Yelina, N.E., Ziolkowski, P.A., Miller, N., Zhao, X., Kelly, K.A., Muñoz, D.F., Mann, D.J., Copenhaver, G.P. and Henderson, I.R. (2013) High-throughput analysis of meiotic crossover frequency and interference via flow cytometry of fluorescent pollen in Arabidopsis thaliana. *Nat. Protoc.* **8**, 2119–2134.
- Yelina, N.E., Lambing, C., Hardcastle, T.J., Zhao, X., Santos, B. and Henderson, I.R. (2015) DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in Arabidopsis. *Genes Dev.* **29**, 2183–2202.
- Ziolkowski, P.A., Berchowitz, L.E., Lambing, C. *et al.* (2015) Juxtaposition of heterozygosity and homozygosity during meiosis causes reciprocal crossover remodeling via interference. *Elife*, **4**, e03708.
- Ziolkowski, P.A., Underwood, C.J., Lambing, C. *et al.* (2017) Natural variation and dosage of the HEI10 meiotic E3 ligase control Arabidopsis crossover recombination. *Genes Dev.* **31**, 306–317.