2516–2528 *Nucleic Acids Research, 2023, Vol. 51, No. 6* *Published online 19 January 2023*
https://doi.org/10.1093/nar/gkac1244

NAR Breakthrough Article

The FANCC–FANCE–FANCF complex is evolutionarily conserved and regulates meiotic recombination

Dipesh Kumar Singh[1], Rigel Salinas Gamboa[1], Avinash Kumar Singh[2], Birgit Walkemeier[1], Jelle Van Leene[3,4], Geert De Jaeger[3,4], Imran Siddiqi[2], Raphael Guerois[5], Wayne Crismani[6,7] and Raphael Mercier[1,*]

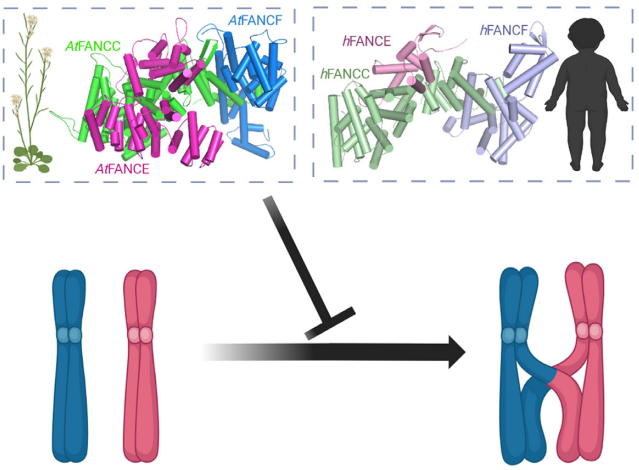
[1]Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, 50829 Cologne, Germany, [2]CSIR-Centre for Cellular & Molecular Biology, Uppal Road, Hyderabad 500007, India, [3]Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent B-9052, Belgium, [4]Center for Plant Systems Biology, VIB, Ghent B-9052, Belgium, [5]Institute for Integrative Biology of the Cell (I2BC), Commissariat à l'Energie Atomique, CNRS, Université Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette 91190, France, [6]The DNA Repair and Recombination Laboratory, St Vincent's Institute of Medical Research, Melbourne 3065, Australia and [7]The Faculty of Medicine, Dentistry and Health Science, The University of Melbourne, Parkville, Victoria, Australia

Received July 07, 2022; Revised November 29, 2022; Editorial Decision December 01, 2022; Accepted December 14, 2022

ABSTRACT

At meiosis, programmed meiotic DNA double-strand breaks are repaired *via* homologous recombination, resulting in crossovers (COs). From a large excess of DNA double-strand breaks that are formed, only a small proportion gets converted into COs because of active mechanisms that restrict CO formation. The Fanconi anemia (FA) complex proteins AtFANCM, MHF1 and MHF2 were previously identified in a genetic screen as anti-CO factors that function during meiosis in *Arabidopsis thaliana*. Here, pursuing the same screen, we identify FANCC as a new anti-CO gene. FANCC was previously only identified in mammals because of low primary sequence conservation. We show that FANCC, and its physical interaction with FANCE–FANCF, is conserved from vertebrates to plants. Further, we show that FANCC, together with its subcomplex partners FANCE and FANCF, regulates meiotic recombination. Mutations of any of these three genes partially rescues CO-defective mutants, which is particularly marked in female meiosis. Functional loss of FANCC, FANCE, or FANCF results in synthetic meiotic catastrophe with the pro-CO factor MUS81. This work reveals that FANCC is conserved outside mammals and has an anti-CO role during meiosis together with FANCE and FANCF.

GRAPHICAL ABSTRACT



INTRODUCTION

Large-scale exchange of genetic material between homologous chromosomes in the form of meiotic crossovers (COs) generates new allelic combinations in the sexual progeny of eukaryotes. COs are also required for the correct segregation of chromosomes at the first meiotic division in most species. This is likely why a mechanism exists to ensure an 'obligate' crossover per chromosome pair, per meiosis. Meiotic recombination is initiated by the formation of a large number of programmed DNA double-stranded breaks

*To whom correspondence should be addressed. Tel: +49 221 5062 401; Email: mercier@mpipz.mpg.de

(DSBs), a minority of which are repaired as COs. Two pathways contribute to CO formation, defining two classes of COs. Class I COs depend on a group of proteins called ZMMs, an acronym derived from seven proteins initially described in *Saccharomyces cerevisiae* (Zip1–Zip4, Msh4–5, Mer3) (1,2), and account for most of the COs. Class II COs, which account for a minority of COs in most eukaryotes including mammals and plants, involve notably the MUS81 nuclease (3,4).

In *Arabidopsis thaliana*, a mutation in any member of the *ZMMs* causes a drastic decrease in CO number, and notably the loss of the obligate crossover, with only a few residual COs formed by the class II pathway, leading to chromosome mis-segregation and quasi-sterility. A forward genetic screen for restoration of fertility of *zmm* mutants identified a series of genes that actively limit CO formation in Arabidopsis. These factors specifically limit class II COs and act through three mechanisms. The first anti-CO pathway involves proteins from the Fanconi anaemia (FA) pathway, FANCM (5), MHF1 and MHF2 (6). These three proteins have been shown to physically interact in humans, along with FAAP24, which has not been identified in Arabidopsis (7,8). The second anti-CO mechanism involves RECQ4, RMI1 and TOP3a (9), and the third, the proteins FIGL1 and FLIP (10–12). These three mechanisms contribute in parallel to limiting class II COs (13).

The FA pathway is comprised of at least 23 protein subunits in human cells and some, but not all, of them are widely conserved in eukaryotes, including plants. It is traditionally known for its role in inter-strand crosslink repair in somatic cells in humans, with emerging roles in replication fork protection (14). The FA pathway is heavily studied because its proper function is required to prevent serious human disease: FA functions as a tumor suppressor and mutation of FA pathway factors causes the rare condition Fanconi anemia. FA pathway proteins can be classified into three groups on the basis of their molecular roles: (i) the FA core complex localizes to DNA inter-strand crosslink or stalled replication forks, with FANCM acting as an anchor. MHF1–MHF2, a heterotetrameric protein complex, promotes FANCM recruitment at the site of DNA damage (7,15). (ii) The FA core complex mono-ubiquitinates each protomer of the FANCI-FANCD2 (ID2) heterodimer. The post-translational modification results in the closing of the ID2 clamp on DNA (16–21). (iii) FA/HR complex proteins are downstream partners that are considered to function independently of the above two groups (22,23).

Structural studies (24–26) have demonstrated that seven subunits of the core complex (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL), and two FA-associated proteins (FAAP20 and FAAP100) form three different subcomplexes: (i) FANCB–FANCL–FAAP100 (BL100), (ii) FANCC–FANCE–FANCF (CEF) and (iii) FANCA–FANCG–FAAP20 (AG20) (26,27). The ring finger domain of FANCL acts as an E3 ubiquitin ligase and its two associated proteins FANCB and FAAP100 are organized as a catalytic module (24,26). It has been proposed that FANCA and FANCG form a chromatin-targeting module, while FANCC, FANCE and FANCF organize to establish a substrate-recognition module (24). FANCF acts as a bridge between FANCC and FANCE (24,28). FANCM interacts with the core complex through FANCF (26,29), demonstrating that the substrate-recognition module is an important component of the FA core complex.

In this study, extending a previously described forward genetic *zmm* suppressor screen (5,6,9–11) augmented by complementary approaches, we demonstrate that the CEF complex is evolutionarily conserved from mammals and show that it is a novel meiotic anti-CO factor.

MATERIALS AND METHODS

Genetic material

The following Arabidopsis lines were used in this study: *fancc-2* (N542341), *fancc-3*(N1007952), *fancc-4* (N626745), *fance* (N553587) (6) *fancf* (N457070) msh4 (N636296)(30), *msh5-2* (N526553) (30), *mus81*-2 (N607515) (31) and *mhf1-3*(N576310) (6). All the T-DNA mutants were obtained from the NASC. Plants were grown under a controlled environment with a photoperiod of 16 h per day and 8 h per night, at a temperature of 20°C, and 70% humidity.

Genetic analysis

The *msh4* suppressor *Atfancc* was sequenced using IIlumina technology at The Genome Analysis Centre, Norwich, UK, and mutations were identified using ler 1 assembly as reported for the MutDetect pipeline (6,32). The identified causal mutation in *fancc-1* was a G to A substitution at position chr3: 21918909 (Ler-0) equivalent to position chr3: 22288888 in the Columbia (TAIR10) genome. The primers used for genotyping are listed in Supplementary Table S1. Siliques were fixed in 70% ethanol for at least two days and scanned for seed counting manually on images. Fertility was assessed by counting seeds per fruit on a minimum of five plants and ten fruits from each plant.

Sequence analyses

Sequences of *A. thaliana* At3g60310/AtFANCC, FANCE (Q9SU89_ARATH) and FANCF (F4K7F0_ARATH) were used as input for the HHpred remote homology detection server against different eukaryotic profile databases (33,34) and as queries of PSI-BLAST searches (35) against the nr database. Full-length sequences of FANCC orthologs were retrieved and re-aligned with mafft (36) and the multiple sequence alignment was represented using JalView (37). The phylogenetic tree of the FANCC orthologs in plants was generated using the FANCC MSA as a query of the PhyML 3.0 server (38) with standard estimated options, an approximate likelihood-ratio test to estimate the bootstrap values (SH-like), and the Jones–Taylor–Thornton substitution model with four substitution rate categories. The calculated tree was represented using the iTOL server (39).

Structural modeling

Sequences of *A. thaliana* At3g60310/AtFANCC, FANCE (Q9SU89_ARATH) and FANCF (F4K7F0_ARATH) were used as input for the MMseqs2 homology search program (40) to generate a multiple sequence alignment (MSA) against the UniRef30 clustered database for each of the

FANC complex subunits (41). The calculated full-length sequences of the orthologs were retrieved and re-aligned with mafft (36). MSAs of FANCC, FANCE and FANCF were then concatenated, matching the sequences of the same species resulting in paired alignments, which were combined with the unpaired sequences for those species that could have not been matched. The resulting paired plus unpaired concatenated MSA was used as input to generate five structural models of the FANCC–FANCE–FANCF complex using a local version of the ColabFold interface (42) running three iterations of the Alphafold2 algorithm (43) trained on the multimer dataset (44) on a local HPC equipped with NVIDIA Ampere A100 80Go GPU cards. The five models converged toward similar conformations and obtained high confidence and quality scores with pLDDTs in the range [79.1, 85] and [72.6, 80.2] and pTMscore in the range [0.64, 0.662]. The model with highest pTMscore was relaxed using rosetta relax protocols to remove steric clashes constrained by the starting structure using the -relax:constrain_relax_to_start_coords option (45), and the model with the lowest rosetta global energy was used for structural analysis. Conservation analysis mapped at the structure of the model were performed using the ConSurf server (46).

Cytological techniques

Meiotic chromosome spreads on anthers were performed as previously described (47). For female meiotic chromosome spreads, floral buds of 0.7–1.4 mm in size (buds with yellow anther) were selected and pistils were isolated. Stigma were removed and the isolated pistil were digested in 0.3% (w/v) cellulase, 0.3% (w/v) pectolyase Y23, 0.3% (w/v) driselase in citrate buffer for 1 h at 37°C. Ovules were isolated from pistils and macerated in minimum amount of water. Further processing was carried out in the same manner as with the male (47). However, the frequency of meiotic cells per slide was substantially lower in females than in males. Chromosomes were stained with DAPI (1 μg/ml) Images were acquired and processed using a ZEISS microscope (AXIO-Imager.Z2) under a 100× oil immersion objective with ZEN software and figures were prepared using Adobe Photoshop.

Yeast two-hybrid and pull down

Clones were generated using the Gateway cloning system (Thermo Fisher Scientific); the desired inserts were cloned into pDONR221 as pENTR clones and then into different destination vectors using the LR clonase recombination method (Thermo Fisher Scientific). We generated full-length ORF pENTR clones for AtFANCC, AtFANCE, and AtMHF2 from an inflorescence cDNA library of Arabidopsis. One additional ORF pENTR of AtFANCC was cloned without a stop codon for in-frame C-terminal fusion and both ORF pENTR clones of AtFANCC were used for GSrhino-tagged pulldown. In yeast two-hybrid assays, we used two destination vectors, pGADT7-GW as bait and pGBKT7-GW as prey. The ORF pENTR clones of AtFANCC, AtFANCE and AtMHF2 were cloned into both destination vectors by LR reaction. The ORF of AtFANCF was cloned into the pGBKT7 and PACT2 AD conventional vector using the NCO1, Sal1, and NCO1, XHO1 restriction enzymes, respectively. All pENTR clones and final clones were verified thoroughly by Sanger sequencing to ensure mutation-free cloning and in-frame fusion. Plasmids of bait and prey clones were transformed into the haploid yeast strains AH109 and Y187, and then yeast two-hybrid assays were performed in a Gal4-based system from Clontech in a diploid strain by mating as previously described (9,48).

Arabidopsis cell suspension cultures expressing N-terminal GSrhino-tagged FANCC and for C-terminal GSrhino-tagged FANCC were used for pull-down as previously described (49,50). Co-purified proteins were identified using standard protocols utilizing on bead-digested sample evaluated on a Q Exactive mass spectrometer (Thermo Fisher Scientific) (51). After identification, the protein list was filtered for false-positives using a list of non-specific proteins, which was assembled as previously reported (51). Semi-quantitative analysis using the average normalized spectral abundance factors (NSAF) of the discovered proteins in the FANCC pull-downs was used to identify true interactors that may have been filtered out due to their classification in the list of nonspecific proteins. Chosen proteins were identified with at least two peptides in at least two experiments and showed high (at least 10-fold) and significant [$\log_{10}(P$ value (t test)) enrichment relative to estimated average NSAF values from a large dataset of pull-downs with nonrelated bait proteins.

FTL analysis

To measure recombination, we used fluorescent transgenic lines (FTL) (420) generated in a Col background. The used lines harbor seed coat expressing GFP (Chr 3:256 516-GFP) and dsRed (Chr 3:5 361 637-dsRed2) fluorescent protein markers in cis (52,53). We quantified the fluorescence of the seeds using the Fiji image analysis software (54), which identifies seeds and quantifies fluorescence intensity for each seed in all pictures. The output was analyzed using a pipeline that was created to normalize the data, plot the frequency of objects with each fluorescent color, plot the fluorescence intensity, and quantify the number of seeds with only one fluorescent color, allowing selection of the number of recombinant seeds. For F2, recombination was measured using the formula below, as reported in (52).

$$cM = 100 * \left(1 - \left[1 - \frac{2(N_G + N_R)}{N_T}\right]^{1/2}\right.$$

For male and female backcrosses, recombination was measured as

$$cM = 100 * \frac{(N_G + N_R)}{N_T}$$

where N_G is the number of green-only fluorescent seeds, N_R is the number of red-only fluorescent seeds and N_T is the total number of seeds counted.

We generated a segregating population (F2) from which we selected plants heterozygous for the markers in cis with

Table 1. Summary of the *zmm* suppressor screen results

		Number of alleles identified in each screen					
	Number of M2 populations screened	total ~7000	*zip4* ~2000	*hei10* ~2000	*msh5* ~1000	*shoc1* ~1000	*msh4 (Ler)* ~1000
identified gene	*FANCC*	**1**					1
	FANCM	**29**	5	5	4	8	7
	MHF2	**4**	1	2			1
	TOP3	**4**		3			1
	RECQ4	**3**					3
	RMI1	**2**				1	1
	FIGL1	**15**	2	5	3	3	2
	FLIP	**1**		1			
	total	**59**	8	16	7	12	16

A suppressor screen was performed in five *zmm* mutants (*zip4, hei10, msh5, shoc1* and *msh4*). One to two thousand independent M2 populations were screened for each *zmm*. At total of 59 causal mutations falling in eight genes were identified and we report the number of alleles identified per gene and per screen. The colors represent three distinct pathways (FA blue, BTR green, FIGL1 red). Note that *msh4* is in the Ler background while others are in Col. As *RECQ4* is duplicated in Col but not in Ler, it could be found only in the *msh4* screen. The list of mutant alleles is shown in Supplementary Table S2.

the desired mutants and wild-type control. For each genotype, we used at least three biological replicates (independent plants) with at least three technical replicates, each of them containing a minimum of 400 seeds. To measure CO frequency independently in males and females, reciprocal crosses were made with wild-type Columbia (0). Differences between genotypes were tested by Chi2 on the proportion of recombined seeds (Ng + Nr) among total seeds.

RESULTS

Identification of a novel *zmm* suppressor

CO-deficient *zmm* mutants display a >90% reduced seed set in Arabidopsis, which correlates with shorter fruit, due to random segregation of chromosomes in meiosis. Therefore, inactivating anti-CO genes in a *zmm* mutant leads to an increase in CO number, resulting in improved chromosome segregation, restored fertility and longer fruits. Here, we extended a forward genetic screen for *zmm* mutants exhibiting increase in fruit length following EMS mutagenesis. The screen was previously performed on five *zmm* mutants (*hei10, zip4, shoc1, msh5* and *msh4*), in a total of ~7,000 lines and identified 59 mutants with restored fertility, among which 58 are mutated in one of the previously identified anti-CO genes (Table 1, Supplementary Table S2). (5,6,9–11). In this study, we focused on the last *zmm* suppressor mutation that increased the fertility of a *msh4* mutant (cshl_GT14269, L*er* genetic background). Genetic mapping delimited the causal mutation to a 0.47MB region on chromosome 3 (21452882–21919909 in the Ler assembly) (55), and whole genome sequencing identified a candidate mutation in the fourth exon donor splicing site of the At3g60310 gene (G > A 3_21918909 in the Ler assembly, corresponding to 3 22288888 in Col TAIR10). We show below that At3g60310 encodes the Arabidopsis FANCC ortholog. Three independent T-DNA alleles (*fancc-2 N542341*, *fancc-3 N1007952* and *fancc-4 N626745*, Col background) were able to enhance the fertility of *msh4*, from 4.5 to > 24 seeds per fruit (Figure 1 and Supplementary Figures S1, S2), confirming the identification of the causal mutation in At3g60310.

FANCC is conserved in plants

Standard sequence similarity analysis failed to find any homology of the protein encoded by At3g60310 outside of plants, or with proteins of known function (56). Using the HHpred remote homology detection server (33,34), it was possible to identify a potential match with human FANCC (XP_011516668) despite both proteins sharing only 16% primary sequence identity (HHpred probability of 94%). To test the hypothesis that At3g60310 is an ortholog of FANCC, we analyzed the physical contacts between human FANCC and the human FANC complex, the cryo-EM structure of which was recently determined at 3.1 Å (57). Figure 2A illustrates that human FANCC (pale green subunit) is in direct physical contact with three subunits of human FA core complex, FANCE (light pink), FANCF (light blue) together with the ubiquitin E3 ligase FANCL (yellow). Given that FANCC, FANCE, and FANCF are known to constitute a stable sub-complex in humans, we tested the possibility that At3g60310 was a dedicated partner of *A. thaliana* FANCE (6) (Q9SU89_ARATH) and FANCF (F4K7F0_ARATH) using the Alphafold2 prediction method (43). Alphafold2 was recently shown to perform well when predicting structures of proteins and whether two proteins interact with each other (44). Using the AlphaFold2 method trained on multimers (44), we obtained a model of the complex with the three *A. thaliana* subunits with reliability scores above the confidence threshold of 50 and 0.5 for pLDDT and ptmscore, respectively (pLDDT of 72.6 and ptmscore of 0.67) (Supplementary Figure S3). Interestingly, At3g60310 was predicted to form a complex with AtFANCE and AtFANCF with a similar arrangement to that observed for the corresponding orthologs in the human FANC complex (Figure 2B). As a support for the reliability of the model, the surface patches 1 and 2 of At3g60310 (circled in Figure 2C) are involved in the interaction with AtFANCF and AtFANCE respectively and are among the most conserved regions of At3g60310/FANCC (Figure 2C, Supplementary Figure S4). The N-terminal domain of FANCE is found well anchored in the central region of At3g60310/FANCC with low predicted error for the accuracy of the interface modelling (Supplementary Figure S3C). When Al-

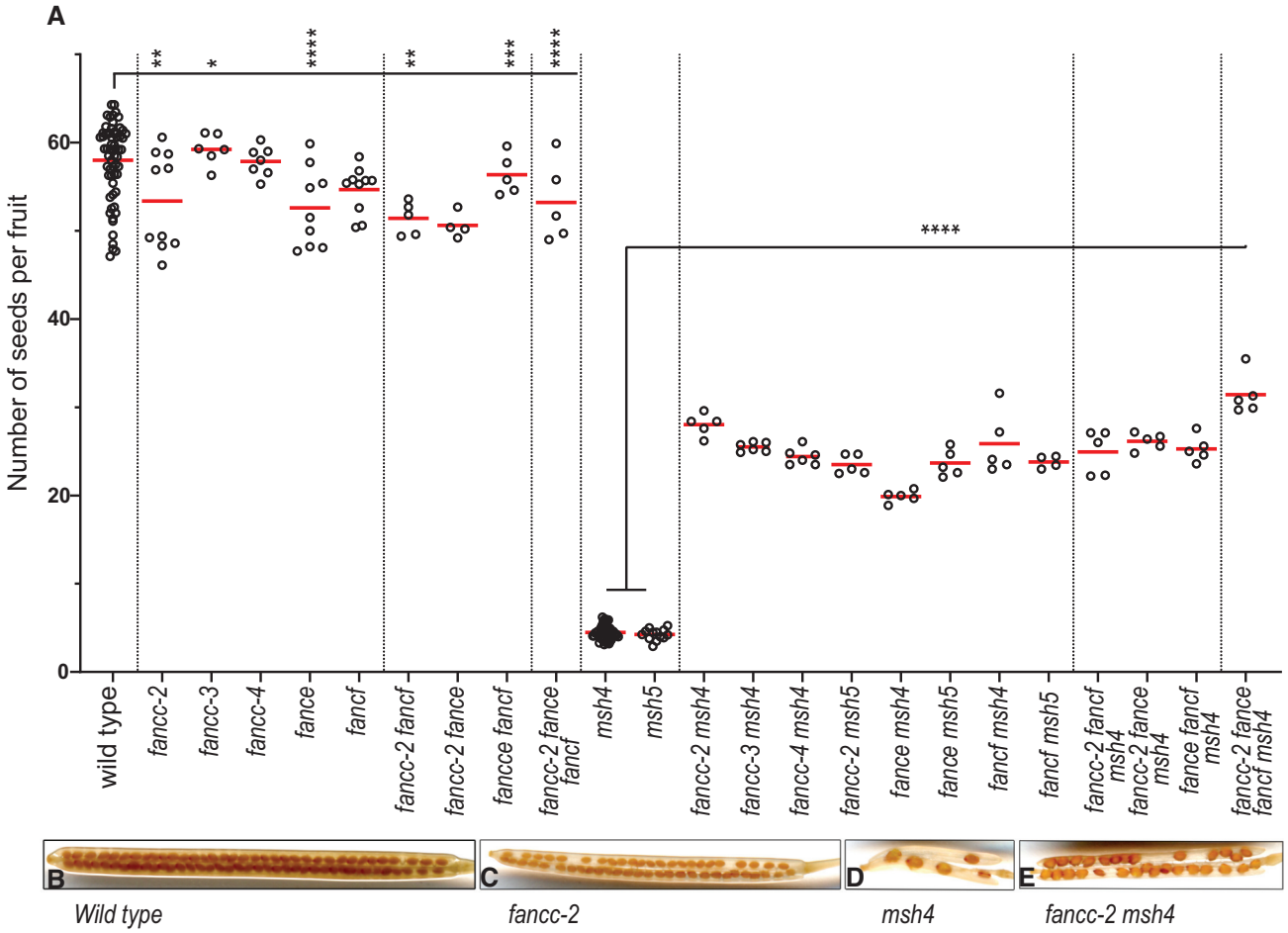


Figure 1. Analysis of fertility of *zmm* suppressor mutants. (**A**) Each dot indicates the fertility of an individual plant, measured as the number of seeds per fruit averaged on 10 fruits. The mean fertility for each genotype is shown by a red bar. Each mutant was compared to sibling controls grown together, and the data of independent experiments are shown in Supplementary Figure S2. Some genotypes were represented in several experiments and their data were pooled for this figure. Stars summarize the one-way ANOVA followed by Sidak test shown in Supplementary Figure S2. (**B**–**E**) Representative fruits of *wild type, fancc-2, msh4*, and *fancc-2 msh4*, cleared with 70% ethanol.

phaFold2 is executed, five models are typically generated whose structural similarity accounts for the robustness of the prediction. In the N-terminal region of FANCE, the five models superimpose very well confirming that the interface between this domain and At3g60310/FANCC is strongly constrained (Supplementary Figure S3B). In contrast, the C-terminal domain of FANCE does not exhibit a strong co-evolutionary signal in the region where it binds to At3g60310/FANCC. The predicted error associated with FANCE C-terminal binding to At3g60310/FANCC is much higher than for the N-terminal domain (Supplementary Figure S3C) and the five models generated for FANCE adopt variable orientations in the C-terminal region (Supplementary Figure S3B). Therefore the relative position of FANCE C-terminal with respect to At3g60310/FANCC should be seen as loosely constrained.

Next, we performed an unbiased search for interacting partners of At3g60310 using pull-down protein purification coupled with mass spectrometry. We used overexpressed GSrhino-tagged At3g60310 as a bait in Arabidopsis suspension cell culture (49). After filtering co-purified proteins for false positives, we recovered peptides from At3g60310 itself and a series of additional proteins in three replicate experiments (Table 2). Strikingly, all four co-purified identified proteins were Arabidopsis homologs of members of the FA complex, FANCE, FANCL, FANCM and MHF2 (Table 2). Further, a yeast two-hybrid assay confirmed direct interactions of At3g60310 with FANCE and MHF2. FANCE, FANCF and MHF2 also interacted with each other in yeast two-hybrid (Supplementary Figure S5). Altogether, this demonstrates that At3g60310 encodes the FANCC protein in Arabidopsis, which we term AtFANCC.

FANCC is conserved in distant eukaryotic lineages

Using PSI-BLAST searches, AtFANCC orthologs could be detected in most plants (Supplementary Figures S4 and S6). In-depth analyses using either PSI-BLAST or HHpred failed to detect any homolog in more distantly related green algae such as *Chlamydomonas*, although a FANCE homolog can be detected in *Chlamydomonas reinhardtii*. In metazoans, a previous bioinformatics analysis performed on model species for all the genes of the FANC core complex, noted that several species were missing a FANCC

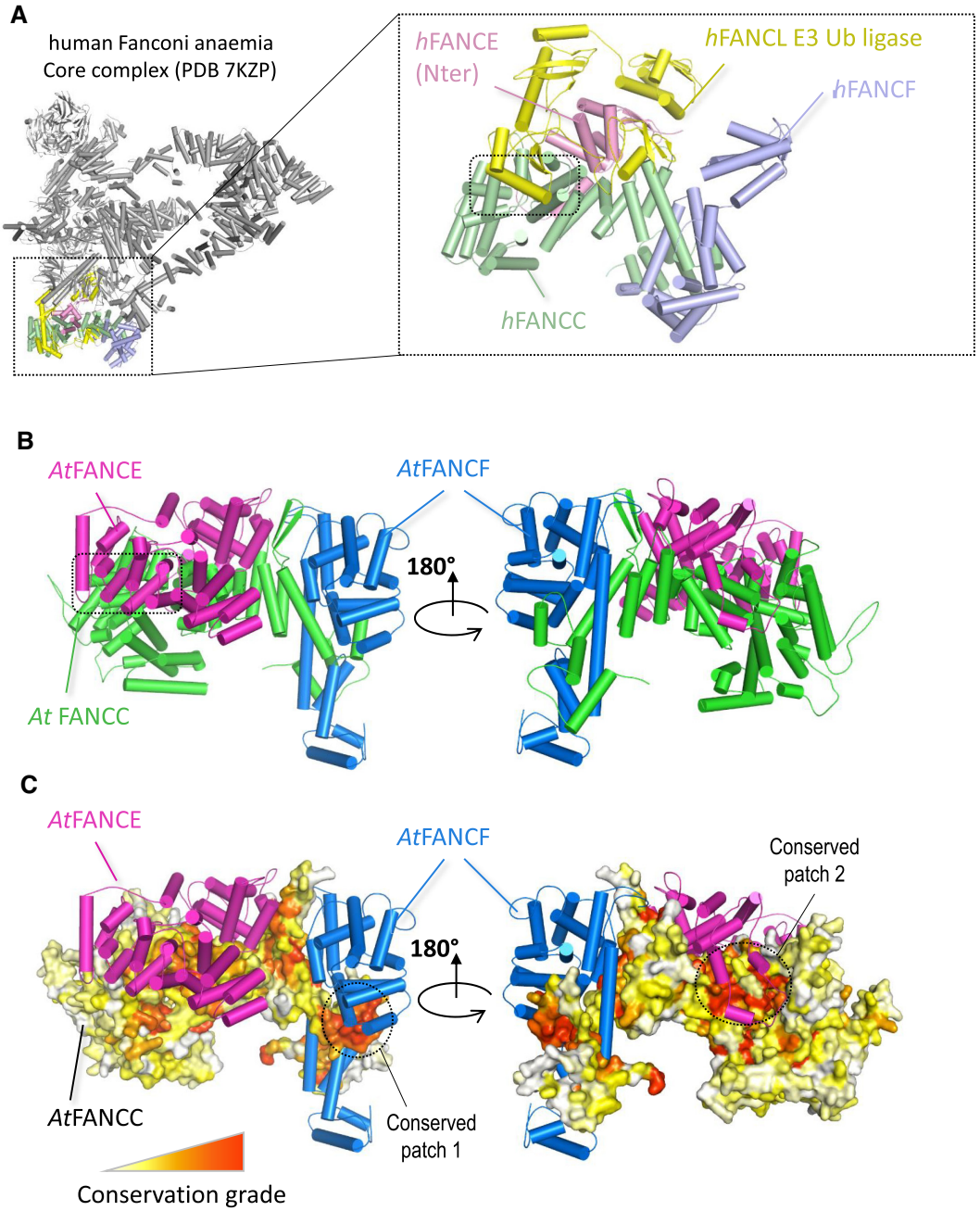


Figure 2. Structural analysis of the experimental human and modeled A. *thaliana* FANCC-FANCE–FANCF complexes. (**A**) Structural representation of the human FANC core complex (PDB:7KZP) (57) Most of the subunits are shown in gray with the exception of those in direct contact with the human FANCC (light green), namely hFANCE (light pink), hFANCL (yellow) and hFANCF (light blue). A zoomed-in view of the four subunits is shown in the inset on the right with the contact region between hFANCL and hFANCC highlighted by a dotted rectangle. (**B**) AlphaFold2 structural model of the AtFANCC-AtFANCE-AtFANCF complex represented as a cartoon in two orientations with a dotted square indicating the C-terminal domain of FANCE located in a region of FANCC that directly binds to the FANCL subunit in the human FANC core complex structure. (**C**) Same view as (B) with AtFANCC shown as a surface and colored according to conservation from white to red for the least to most conserved positions. Pymol software was used to draw the different structures (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

homolog although having a FANCE ortholog (56). We revisited this study using the most recent sequence databases and PSI-BLAST searches starting from human FANCC. Interestingly, five iterations of PSI-BLAST were required to retrieve the first plant ortholog (in the monocot *Spirodela intermedia*), which enabled the retrieval of all the same plant orthologs identified from AtFANCC. After 15 iterations, the search nearly converged with about 2,100 FANCC homologs, highlighting the existence of FANCCs in early branching metazoans such as *Nematostella vectensis* (XP_032241565) and *Ciona intestinalis* (XP_002129616) that were not found previously. In insects, orthologs could also be detected in Hymenoptera (ants and bees) but neither in Diptera (Drosophila) nor in Lepidoptera (Bombyx).

Table 2. Pull-down protein purification using AT3G60310/FANCC as bait

Gene Id	Protein name	CGSrhino PD1	CGSrhino PD2	CGSrhino PD3
AT3G60310	FANCC	19	18	15
AT4G29560	FANCE	6	4	5
AT5G65740	FANCL	5	4	4
AT1G78790	MHF2	2	-	3
AT1G35530	FANCM*	2	-	-

Three replicates of pull-down purifications (PD1, PD2 and PD3) followed by mass spectrometry were performed using FANCC as a bait. After filtering (see material and methods), the number of specific peptides is reported for each identified protein. *Only identified in one experiment. MS data are shown in Supplementary Table S4.

Consistently, repeating the PSI-BLAST search with human FANCE or FANCF as queries, a similar distribution of homologs was found in insects. Homologs of FANCE and FANCC could be detected in specific fungal lineages such as *Rhizopus azygosporus* (corresponding to hypothetical proteins RCH90546.1 and RCH79564.1, respectively). A reciprocal HHpred analysis comparing these genes against the human database confirmed they were remote homologs of FANCE and FANCC (HHpred probability score of 100% and sequence identities of 21% and 15%, respectively), suggesting that certain fungal lineages did not lose these FANC complex subunits.

Atfancc, atfance, and *atfancf* mutations increase fertility and bivalent formation of crossover-deficient *zmm* mutants

As the mammalian FANCC was shown to form a structural and functional module with FANCE and FANCF (24,25,57), we explored their potential meiotic roles in parallel. A FANCE homolog was previously described in Arabidopsis (AT4G29560) (6), and the AT5G44010 gene was annotated as *AtFANCF* in Araport11 because of sequence similarity with the mammalian *FANCF* (58). The Arabidopsis *FANCC*, *FANCE* and *FANCF* genes are expressed both in somatic and reproductive tissues (59). Single mutants in each of these genes did not show growth or developmental defects but had a slight decrease in fertility (Figure 1A–C, Supplementary Figure S2). Meiotic chromosome spreads in *Atfance* and *Atfancf* single mutants revealed the presence of univalents at low frequencies, showing that some chromosome pairs lack COs (Supplementary Figure S7, Figure 3C–D, G). Combination of *fancc, fance* and *fancf* mutations did not reveal any developmental defects or enhanced sterility and meiotic defects (Figures 1A and 3G). This suggests that, consistent with the pull-down and Y2H data, FANCC, FANCE and FANCF act together at meiosis, playing a role in ensuring the obligate CO.

Mutation of *FANCC*, or *FANCE*, or *FANCF* significantly restored the fertility of the *zmm* mutants *msh4* and *msh5*, increasing the seed set more than fourfold (Sidak test, $P < 10^{-6}$) (Figure 1A, D, E. Supplementary Figure S2). In the Ler background, chromosome spreads of male meiosis showed an increase of bivalent frequency in *fancc-1 msh4* compared to *msh4* ($P < 0.001$) (Figure 3G). In comparable experiments in the Col background, *fancc*, *fance* or *fancf* individual mutations barely increased bivalents in *msh4* ($P = 0.12$, 0.12 and 0.0002, respectively) (Figure 3G). This mild effect could explain why the anti-CO effect of *FANCE* was missed in previous studies (6). Combining Col *fancc, fance* and *fancf* mutations in *msh4* further restored bivalent formation to reach an average of 2.8 bivalents/cell compared to 1.4 in *msh4* ($P < 0.0001$) and led to a slightly higher fertility increase compared to *fancc msh4* ($P = 0.04$). Altogether, this suggests that FANCC, FANCE and FANCF limit CO formation in a partially redundant manner. Note that this restoration of bivalent formation is lesser than that obtained through mutation of *FANCM* (5 bivalents) (5), MHF1 or MHF2 (3.6 bivalents) (Figure 3G) (6), suggesting that the FANCC–FANCE–FANCF module has a supporting role in limiting a portion of the COs prevented by FANCM-MHF1-MHF2. The *fancc* mutation did not further restore bivalent formation in *mhf1 msh4* (*t*-test $P = 0.56$), suggesting that FANCC acts in the same anti-CO pathway as MHF1 (Figure 3G).

FANCC, FANCE and FANCF regulate meiotic crossover formation

Intriguingly, in the above experiments, we found that *fancc*, *fance* and *fancf* increased the fertility of *zmm* mutants (*msh4* and *msh5*), but that the increase of bivalent number in male meiotic cells was less robust than the seed set suggested. As fertility in the self-pollinating plant Arabidopsis depends on both male and female meiosis, this may suggest that the role of FANCC, FANCE and FANCF in limiting COs is more critical in female meiosis than in male meiosis. Supporting this suggestion, the average number of bivalents at metaphase I of female meiosis was increased from 1.3 in *msh4* to 3.7 in *fancc-1 msh4* ($n = 9$ and $n = 7$, respectively, *t*-test $P = 0.0005$. Figure 4). Further, we used a test line for recombination (FTL420), which contains two transgenes conferring expression of GFP and RFP in the seed coat and defining a 5-Mb interval of the sub-telomeric region on chromosome 3 (52,53). Crossover frequency was measured for males and females separately, through reciprocal crosses with wild-type plants, and in selfing (Figure 5, Supplementary Table S3). In females, recombination was significantly increased (Fisher test, $P < 0.0001$) in *fancc* and *fancc fance fancf* compared to wild type, confirming the anti-CO function of *FANCC* (Figure 5A). In males, crossover frequency was not increased, but slightly reduced ($P = 0.17$ for *fancc* and 0.015 for *fancc fance fancf*). In selfing, which combines male and female meiosis products, recombination was modestly increased in *fancc* compared to wild type ($P = 0.003$). A similar recombination picture was observed in *fance*, *fancf* and the triple mutant *fancc fance fancf*, suggesting that the three proteins act together in limiting meiotic crossovers (Figure 5A).

fancc, fance, and *fancf* exhibit chromosome fragmentation in the *mus81* background

Because of the roles of FANCM, MHF1 and MHF2 in preventing class II COs, combining mutation in these genes with mutation of *MUS81* that catalyses class II COs, leads to chromosome fragmentation at meiosis, resulting in sterility. In addition, the *fancm mus81* double mutant shows

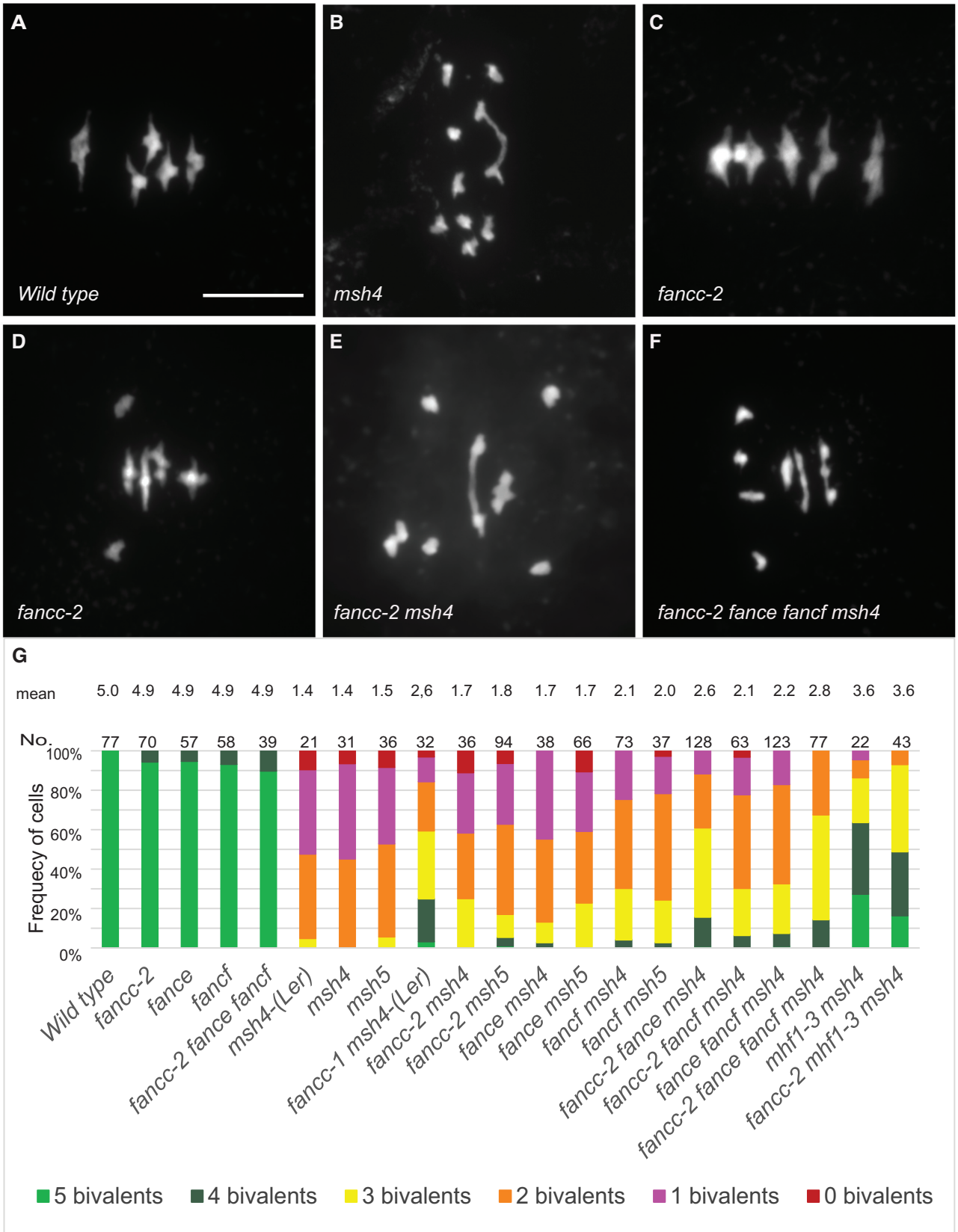


Figure 3. Metaphase I chromosome spreads of male meiocytes. (**A**) Wild type with five bivalents, (**B**) *msh4* with one bivalent, (**C**) *fancc-2* with four bivalents and one pair of univalents. (**D**) *fancc-2* with five bivalents. (**E**) *fancc-2 msh4* with two bivalents. (**F**) *fancc-2 fance fancf msh4* with two bivalents. Scale bar, 10 μm. (**G**) Quantification of bivalents at metaphase I. The proportion of cells with 0–5 bivalents is shown with a color code. The number of analyzed cells and the average bivalent number per cell is shown for each genotype. All the genotypes are in the Col background, except when Ler is mentioned.

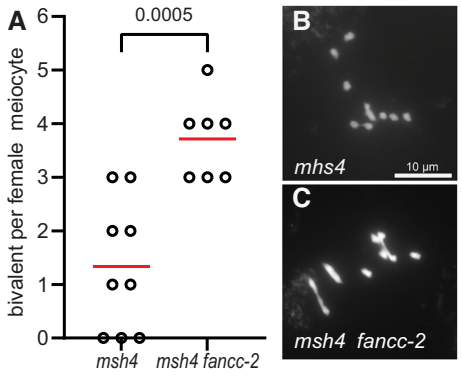


Figure 4. Metaphase I chromosome spreads of female meiocytes. (**A**) Quantification of bivalents at female metaphase I. Each dot indicates the bivalent number of individual meiotic cells. The mean bivalent number for each genotype is shown by a red bar. (**B**) Female *msh4* meiocyte with one bivalent. (**C**) Female *fancc-2 msh4* meiocyte with four bivalents. Scale bar, 10 μm.

a strong developmental defect, demonstrating the role of these two genes in somatic DNA repair (5,6) When we combined *fancc*, *fance*, or *fancf* with the *mus81* mutation, we did not observe developmental defects. However, in double mutants with *mus81* and either *fanc-c*, *-e* or *-f* we observed a strong reduction in fertility, measured by seed per fruit, compared to the respective single mutants (Figure 6A, Supplementary Figure S8). Meiotic chromosome spreads revealed the presence of chromosome fragments at anaphase I and subsequent stages in ∼40% of the cells of the double mutants (Figure 6B–J, Supplementary Figure S9 A–D and F–I). This demonstrates that FANCC, FANCE, and FANCF are important for efficient DSB repair in a *mus81* background and suggests that they regulate class II CO formation but with a less critical role than FANCM and MHF1/2. The removal of all four genes—*mus81 fancc fance fancf*—did not drastically enhanced fertility defects or chromosome fragmentation compared to the double *mus81 fancc* combinations. These results support the hypothesis that all three genes, *FANCC, FANCE* and *FANCF*, are required at meiosis to repair a subset of intermediates that can also be repaired by MUS81.

DISCUSSION

FANCC–FANCE–FANCF constitute a stable subcomplex within the FA core complex. Based on sequence conservation, FANCE and FANCF homologs have been identified in evolutionarily distant eukaryotes such as plants (6,58). However, despite multiple studies that systematically catalogued FA pathway protein conservation across diverse taxa, homologs of FANCC have not been identified beyond vertebrates (6,60), suggesting that FANCC may not be conserved over large evolutionary scales. Here, combining genetics, *in vivo* pull-downs, direct protein-protein interaction studies, and structural modeling, we unambiguously identified the FANCC protein in Arabidopsis. In addition, interaction and modeling studies strongly suggest that FANCC, FANCE and FANCF form a sub-complex in Arabidopsis as they do in vertebrates. Homologs of FANCC can also be readily identified in most

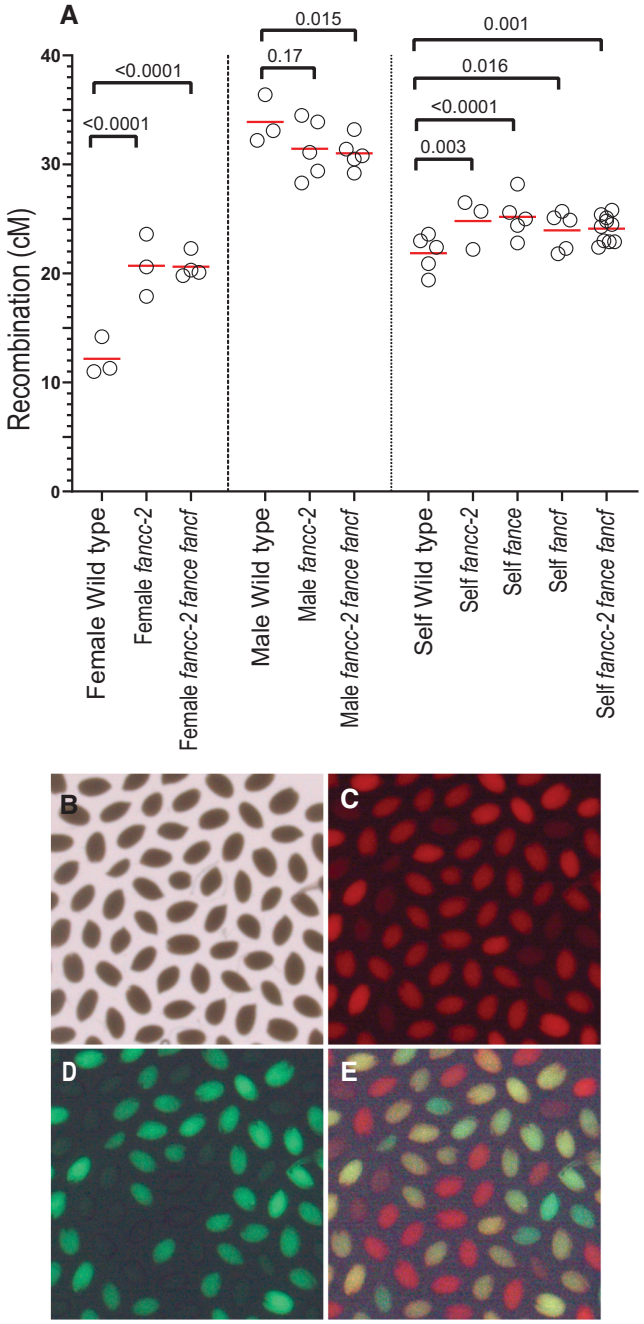


Figure 5. Recombination in *fancc*, *fance* and *fancf* mutant. (**A**) Recombination was measured in seeds produced by crosses with wild type (female and male) or after selfing. Each dot represents the recombination frequency measured in an individual plant, and the red lines show the mean. *P* values are from two-sided Fisher's exact test on the proportion of recombined seeds. Raw data are shown in Supplementary Table S3 (**B**–**E**). Representative image of seeds from a 420/++ hemizygote imaged under bright-field, red fluorescence channel, green fluorescence channel, and merged fluorescence.

other plants. As the plant and animal branches diverged very early in the eukaryotic tree of life (61), this suggests that the FA complex and notably the FANCC–E–F subcomplex was already present in the common ancestor of all living eukaryotes. The algorithm we used to detect

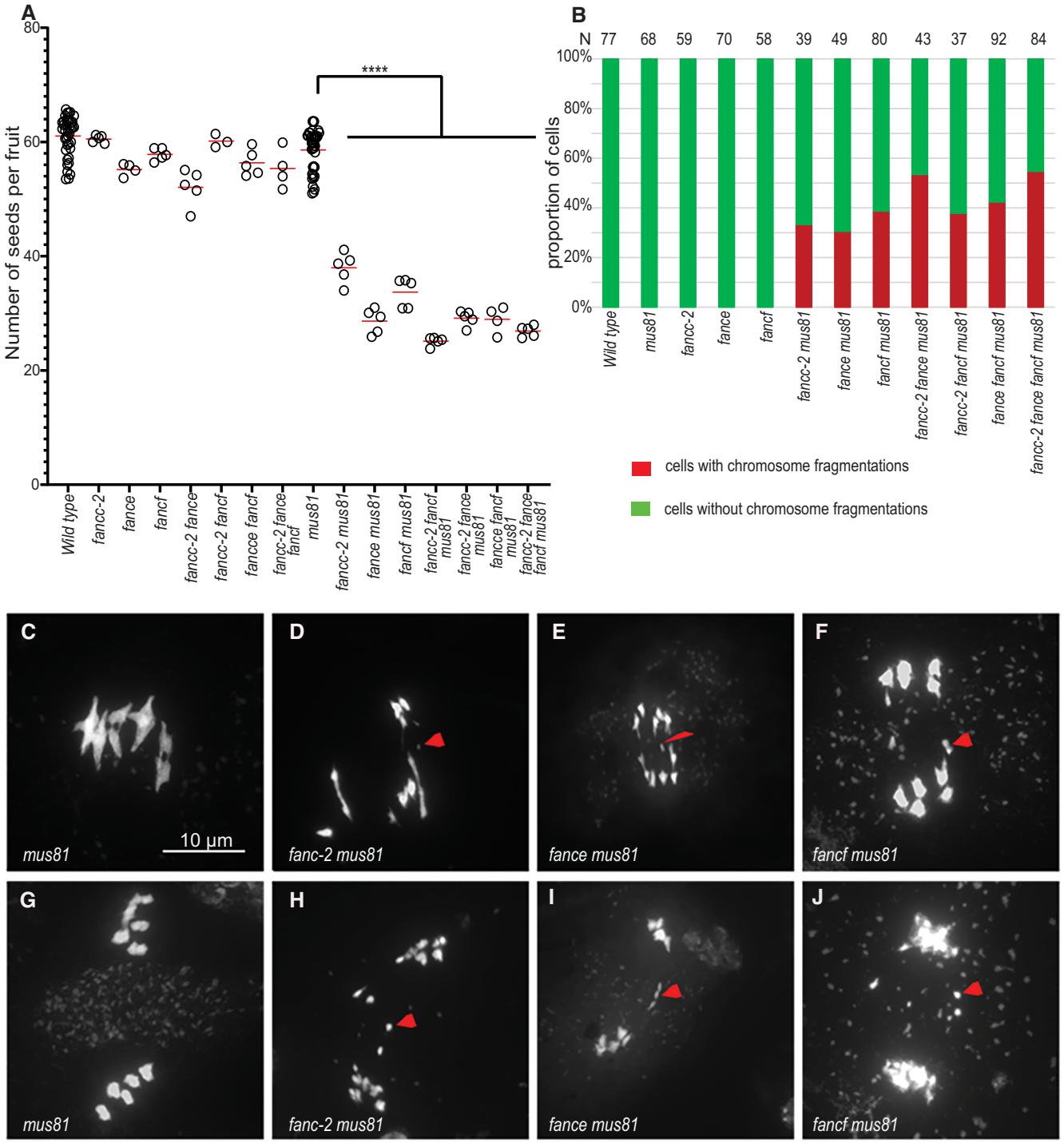


Figure 6. Combining *fanc* and *mus81* mutations leads to reduced fertility and chromosome fragmentation at meiosis. (**A**) Each dot indicates the fertility of an individual plant, measured as the number of seeds per fruit averaged on ten fruits. The means for each genotype are shown by red bars. Each double mutant was compared to sibling controls grown together; the independent experiments are shown in Supplementary Figure S8. Wild type and *mus81* control were represented in several experiments and their data are pooled in this plot. (**B**) Quantification of cells with and without chromosome fragments. N = number of cells analyzed for each genotype. Stars summarize the one-way ANOVA followed by Sidak test shown in Supplementary Figure S8. (C–J) Chromosome spreads of male meiocytes (Scale bar, 10 µm). Arrow heads indicate chromosome fragments. (**C**, **G**) *mus81*. (**D**, **H**) *fancc-2 mus81*. (**E**, **I**) *fance mus81* (**F**, **J**) *fancf mus81*.

divergent homologs succeeded in detecting the link between plant and vertebrate FANCC but failed to detect homologs in fungal lineages, except for a few species. As fungi are more closely related to animals than plants, this suggests that most of the fungal lineages have lost FANCC, or that the FANCC sequence has diverged beyond what we can recognize with current tools. Similarly, FANCC was detected in diverse animal lineages including some insects, but not in Drosophila, which can be attributed either to gene loss or to extreme divergence.

We initially identified *FANCC* because its mutation can partially restore the fertility of CO-defective *zmm* mutants, in a similar manner to previously identified anti-CO factors and notably the FA complex components FANCM, MHF1, and MHF2 (5,6). We also found that mutation in either of the two other subunits of the FA CEF subcomplex, *fance* and *fancf*, improves the fertility of *zmm* mutants. Mutations in the three genes individually restored fertility of *zmm* to similar levels, but to a much lower level than previously obtained with *mhf1*, *mhf2* or *fancm*. Further, restoration of *zmm* fertility upon cumulative mutations in *fancc, fance* and *fancf* remained limited. This suggests that FANCC, FANCE and FANCF together regulate meiotic recombination, but with a less critical role than FANCM, MHF1, and MHF2. We observed an increased number of bivalents at male meiosis when mutating *fancc*, *fance* and *fancf* in *zmm* mutants, consistent with an anti-CO function. However, the increase in bivalents in males was limited compared to the observed increase in fertility, suggesting that male and female meiosis could be differently affected. Indeed, chromosome spreads of female meiosis revealed a large increase of bivalents numbers in *msh4 fancc* compared to *msh4*. We also observed a slight decrease in fertility and a low frequency of univalents in male meiocytes in single *fancc*, *fance* or *fancf* mutants, suggesting a pro-CO function. However, when assessing recombination by a genetic assay in *fancc* and *fancc fance fancf*, we observed a large increase in recombination in females and a small decrease in males. Altogether, we propose that FANCC–E–F regulates meiotic recombination, with a predominant anti-CO function in females, explaining the capacity of their mutation to restore the fertility of *zmm* mutants.

Similar to FANCM and MHF1/MH2, we propose that FANCC–E–F prevents the formation of class II COs that are catalyzed by MUS81. Indeed, combining any of *fancc*, *fance* or *fancf* with the *mus81* mutation led to chromosome fragmentation at meiosis and reduced fertility (Figure 6). Combining the three mutations (*fancc fance fancf*) together had only a slightly increased effect compared to single mutants in the capacity to increase fertility and bivalents of *zmm* mutants (Figures 1 and 3) or for synthetic meiotic catastrophe and reduced fertility when combined with *mus81* (Figure 6). Further, the recombination assay did not detect differences between the single *fancc* and the triple *fancc fance fancf* (Figure 5), suggesting that FANCC–FANCE–FANCF act together in regulating recombination. The capacity of the *fancc fance fancf* mutations to restore bivalent and fertility of the *zmm* mutants and to increase crossover frequencies is weaker than observed with mutation of MHF1/2 or FANCM (5,6). Further, the *fancc-2* has no additive effect with *mhf1* (Figure 3). Altogether, this shows that FANCC–FANCE–FANCF acts in the same anti-CO pathway as FANCM-MHF1/2, as also supported by the fact that they form a stable complex *in vivo* (Table 2). We propose that FANCC–FANCE–FANCF supports FANCM activity, which unwinds recombination intermediates and directs them to non-crossover repair. We favor the hypothesis that the meiotic crossover-limiting role of Arabidopsis FANCC–E–F is distinct from the well-described somatic role of human FANC–C–E–F where it facilitates FANCD2–FANCI mono-ubiquitination in inter-strand crosslink repair. It is unclear if the crossover-limiting role can be uncouple from the canonical role of the FA core complex as no other core complex components appear to exist in Arabidopsis, which would permit assessing this question. This hypothesis is supported by the following lines of evidence: (i) there is no detectable crossover-limiting role of neither of the Arabidopsis orthologues of the catalytic component of the human FA core complex—the E3 RING ligase, FANCL—and its substrate FANCD2-FANCI (6,62); (ii) human FANCM has well-described functions distinct from the FA core complex (29,63,64) that are associated with remodeling branched DNA structures. The FANCC–E–F complex may act to stabilize or support the activity of FANCM in performing its function of branched molecule dissolution during meiotic DSB repair.

Finally, this work shows the value of model organisms to advance knowledge and stimulate new research questions in non-model species such as humans. While *Arabidopsis thaliana* may be a long way from humans, we feel these findings and others (60) create a compelling case that DNA repair functions of FANC genes predate vertebrates, and that there have likely multiple types of DNA repair, both somatic and meiotic.

DATA AVAILABILITY

The data underlying this article are available in the article and in its online supplementary material.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Virginie Portemer for her help in *fancc-1* mapping; Piotr A. Ziolkowski for kindly providing the 420 FTL line; Peter Huijser and Miguel Vallebueno-Estrada for supporting with the FTL image analysis setup; VIB proteomics core facility for performing the Q Exactive analysis of the pull down samples and Neysan Donnelly for proofreading the manuscript.

FUNDING

Max Planck Society (to R.M.); Depart of Biotechnology Centre of Excellence grant (to I.S.); CEFIPRA project SMOKI (to R.M. and I.S.). Funding for open access charge: Max Planck Society.

Conflict of interest statement. None declared.

REFERENCES

1. Borner,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.
2. Pyatnitskaya,A., Borde,V. and De Muyt,A. (2019) Crossing and zipping: molecular duties of the ZMM proteins in meiosis. *Chromosoma*, **128**, 181–198.
3. Youds,J.L. and Boulton,S.J. (2011) The choice in meiosis - defining the factors that influence crossover or non-crossover formation. *J. Cell Sci.*, **124**, 501–513.
4. de los Santos,T., Hunter,N., Lee,C., Larkin,B., Loidl,J. and Hollingsworth,N.M. (2003) The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast. *Genetics*, **164**, 81–94.
5. 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.
6. Girard,C., Crismani,W., Froger,N., Mazel,J., Lemhemdi,A., Horlow,C. and Mercier,R. (2014) FANCM-associated proteins MHF1 and MHF2, but not the other Fanconi anemia factors, limit meiotic crossovers. *Nucleic Acids Res.*, **42**, 9087–9095.
7. Singh,T.R., Saro,D., Ali,A.M., Zheng,X.F., Du,C.H., Killen,M.W., Sachpatzidis,A., Wahengbam,K., Pierce,A.J., Xiong,Y. *et al.* (2010) MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. *Mol. Cell*, **37**, 879–886.
8. Tao,Y., Jin,C., Li,X., Qi,S., Chu,L., Niu,L., Yao,X. and Teng,M. (2012) The structure of the FANCM-MHF complex reveals physical features for functional assembly. *Nat. Commun.*, **3**, 782.
9. Seguela-Arnaud,M., Choinard,S., Larcheveque,C., Girard,C., Froger,N., Crismani,W. and Mercier,R. (2017) RMI1 and TOP3alpha limit meiotic CO formation through their C-terminal domains. *Nucleic Acids Res.*, **45**, 1860–1871.
10. Girard,C., Chelysheva,L., Choinard,S., Froger,N., Macaisne,N., Lemhemdi,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.
11. Fernandes,J.B., Duhamel,M., Seguela-Arnaud,M., Froger,N., Girard,C., Choinard,S., Solier,V., De Winne,N., De Jaeger,G., Gevaert,K. *et al.* (2018) FIGL1 and its novel partner FLIP form a conserved complex that regulates homologous recombination. *PLoS Genet.*, **14**, e1007317.
12. Kumar,R., Duhamel,M., Coutant,E., Ben-Nahia,E. and Mercier,R. (2019) Antagonism between BRCA2 and FIGL1 regulates homologous recombination. *Nucleic Acids Res.*, **47**, 5170–5180.
13. Fernandes,J.B., Seguela-Arnaud,M., Larcheveque,C., Lloyd,A.H. and Mercier,R. (2018) Unleashing meiotic crossovers in hybrid plants. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 2431–2436.
14. Kolinjivadi,A.M., Crismani,W. and Ngeow,J. (2020) Emerging functions of Fanconi anemia genes in replication fork protection pathways. *Hum. Mol. Genet.*, **29**, R158–R164.
15. Yan,Z., Delannoy,M., Ling,C., Daee,D., Osman,F., Muniandy,P.A., Shen,X., Oostra,A.B., Du,H., Steltenpool,J. *et al.* (2010) A histone-fold complex and FANCM form a conserved DNA-remodeling complex to maintain genome stability. *Mol. Cell*, **37**, 865–878.
16. Smogorzewska,A., Matsuoka,S., Vinciguerra,P., McDonald,E.R. 3rd, Hurov,K.E., Luo,J., Ballif,B.A., Gygi,S.P., Hofmann,K., D'Andrea,A.D. *et al.* (2007) Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell*, **129**, 289–301.
17. Wang,R., Wang,S., Dhar,A., Peralta,C. and Pavletich,N.P. (2020) DNA clamp function of the monoubiquitinated Fanconi anaemia ID complex. *Nature*, **580**, 278–282.
18. Alcon,P., Shakeel,S., Chen,Z.A., Rappsilber,J., Patel,K.J. and Passmore,L.A. (2020) FANCD2-FANCI is a clamp stabilized on DNA by monoubiquitination of FANCD2 during DNA repair. *Nat. Struct. Mol. Biol.*, **27**, 240–248.
19. Tan,W., van Twest,S., Leis,A., Bythell-Douglas,R., Murphy,V.J., Sharp,M., Parker,M.W., Crismani,W. and Deans,A.J. (2020) Monoubiquitination by the human Fanconi anemia core complex clamps FANCI:FANCD2 on DNA in filamentous arrays. *Elife*, **9**, e54128.
20. Rennie,M.L., Lemonidis,K., Arkinson,C., Chaugule,V.K., Clarke,M., Streetley,J., Spagnolo,L. and Walden,H. (2020) Differential functions of FANCI and FANCD2 ubiquitination stabilize ID2 complex on DNA. *EMBO Rep.*, **21**, e50133.
21. Lopez-Martinez,D., Kupculak,M., Yang,D., Yoshikawa,Y., Liang,C.C., Wu,R., Gygi,S.P. and Cohn,M.A. (2019) Phosphorylation of FANCD2 Inhibits the FANCD2/FANCI complex and suppresses the Fanconi anemia pathway in the absence of DNA damage. *Cell Rep.*, **27**, 2990–3005.
22. Kottemann,M.C. and Smogorzewska,A. (2013) Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature*, **493**, 356–363.
23. Deakyne,J.S. and Mazin,A.V. (2011) Fanconi anemia: at the crossroads of DNA repair. *Biochemistry (Mosc)*, **76**, 36–48.
24. Shakeel,S., Rajendra,E., Alcon,P., O'Reilly,F., Chorev,D.S., Maslen,S., Degliesposti,G., Russo,C.J., He,S., Hill,C.H. *et al.* (2019) Structure of the Fanconi anaemia monoubiquitin ligase complex. *Nature*, **575**, 234–237.
25. Swuec,P., Renault,L., Borg,A., Shah,F., Murphy,V.J., van Twest,S., Snijders,A.P., Deans,A.J. and Costa,A. (2017) The FA core complex contains a homo-dimeric catalytic module for the symmetric mono-ubiquitination of FANCI-FANCD2. *Cell Rep.*, **18**, 611–623.
26. Huang,Y., Leung,J.W., Lowery,M., Matsushita,N., Wang,Y., Shen,X., Huong,D., Takata,M., Chen,J. and Li,L. (2014) Modularized functions of the Fanconi anemia core complex. *Cell Rep.*, **7**, 1849–1857.
27. Rajendra,E., Oestergaard,V.H., Langevin,F., Wang,M., Dornan,G.L., Patel,K.J. and Passmore,L.A. (2014) The genetic and biochemical basis of FANCD2 monoubiquitination. *Mol. Cell*, **54**, 858–869.
28. Leveille,F., Blom,E., Medhurst,A.L., Bier,P., Laghmani el,H., Johnson,M., Rooimans,M.A., Sobeck,A., Waisfisz,Q., Arwert,F. *et al.* (2004) The Fanconi anemia gene product FANCF is a flexible adaptor protein. *J. Biol. Chem.*, **279**, 39421–39430.
29. Deans,A.J. and West,S.C. (2009) FANCM connects the genome instability disorders Bloom's syndrome and Fanconi Anemia. *Mol. Cell*, **36**, 943–953.
30. Higgins,J.D., Vignard,J., Mercier,R., Pugh,A.G., Franklin,F.C. and Jones,G.H. (2008) AtMSH5 partners AtMSH4 in the class I meiotic crossover pathway in *Arabidopsis thaliana*, but is not required for synapsis. *Plant J.*, **55**, 28–39.
31. 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.
32. Schneeberger,K., Ossowski,S., Ott,F., Klein,J.D., Wang,X., Lanz,C., Smith,L.M., Cao,J., Fitz,J., Warthmann,N. *et al.* (2011) Reference-guided assembly of four diverse Arabidopsis thaliana genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 10249–10254.
33. Zimmermann,L., Stephens,A., Nam,S.Z., Rau,D., Kubler,J., Lozajic,M., Gabler,F., Soding,J., Lupas,A.N. and Alva,V. (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.*, **430**, 2237–2243.
34. Soding,J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics*, **21**, 951–960.
35. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
36. Katoh,K. and Standley,D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, **30**, 772–780.
37. Waterhouse,A.M., Procter,J.B., Martin,D.M., Clamp,M. and Barton,G.J. (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.
38. Dereeper,A., Guignon,V., Blanc,G., Audic,S., Buffet,S., Chevenet,F., Dufayard,J.F., Guindon,S., Lefort,V., Lescot,M. *et al.* (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.*, **36**, W465–W469.
39. Letunic,I. and Bork,P. (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.*, **49**, W293–W296.

40. Steinegger,M. and Soding,J. (2017) MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.*, **35**, 1026–1028.
41. Mirdita,M., von den Driesch,L., Galiez,C., Martin,M.J., Soding,J. and Steinegger,M. (2017) Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Res.*, **45**, D170–D176.
42. Mirdita,M., Schütze,K., Moriwaki,Y., Heo,L., Ovchinnikov,S. and Steinegger,M. (2022) ColabFold: making protein folding accessible to all. *Nat. Methods*, **19**, 679–682.
43. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Zidek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.
44. Evans,R., O'Neill,M., Pritzel,A., Antropova,N., Senior,A., Green,T., Žídek,A., Bates,R., Blackwell,S., Yim,J. *et al.* (2021) Protein complex prediction with AlphaFold-Multimer. bioRxiv doi: https://doi.org/10.1101/2021.10.04.463034, 10 March 2022, preprint: not peer reviewed.
45. Leman,J.K., Weitzner,B.D., Lewis,S.M., Adolf-Bryfogle,J., Alam,N., Alford,R.F., Aprahamian,M., Baker,D., Barlow,K.A., Barth,P. *et al.* (2020) Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat. Methods*, **17**, 665–680.
46. Ashkenazy,H., Abadi,S., Martz,E., Chay,O., Mayrose,I., Pupko,T. and Ben-Tal,N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic. Acids. Res.*, **44**, W344–W350.
47. Ross,K.J., Fransz,P. and Jones,G.H. (1996) A light microscopic atlas of meiosis in Arabidopsis thaliana. *Chromosome Res.*, **4**, 507–516.
48. Rossignol,P., Collier,S., Bush,M., Shaw,P. and Doonan,J.H. (2007) Arabidopsis POT1A interacts with TERT-V(I8), an N-terminal splicing variant of telomerase. *J. Cell Sci.*, **120**, 3678–3687.
49. Van Leene,J., Han,C., Gadeyne,A., Eeckhout,D., Matthijs,C., Cannoot,B., De Winne,N., Persiau,G., Van De Slijke,E., Van de Cotte,B. *et al.* (2019) Capturing the phosphorylation and protein interaction landscape of the plant TOR kinase. *Nat Plants*, **5**, 316–327.
50. Cromer,L., Jolivet,S., Singh,D.K., Berthier,F., De Winne,N., De Jaeger,G., Komaki,S., Prusicki,M.A., Schnittger,A., Guerois,R. *et al.* (2019) Patronus is the elusive plant securin, preventing chromosome separation by antagonizing separase. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 16018–16027.
51. Van Leene,J., Eeckhout,D., Cannoot,B., De Winne,N., Persiau,G., Van De Slijke,E., Vercruysse,L., Dedecker,M., Verkest,A., Vandepoele,K. *et al.* (2015) An improved toolbox to unravel the plant cellular machinery by tandem affinity purification of Arabidopsis protein complexes. *Nat. Protoc.*, **10**, 169–187.
52. Ziolkowski,P.A., Berchowitz,L.E., Lambing,C., Yelina,N.E., Zhao,X., Kelly,K.A., Choi,K., Ziolkowska,L., June,V., Sanchez-Moran,E. *et al.* (2015) Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover remodelling via interference during Arabidopsis meiosis. *Elife*, **4**, e03708.
53. 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.
54. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.
55. Zapata,L., Ding,J., Willing,E.M., Hartwig,B., Bezdan,D., Jiao,W.B., Patel,V., Velikkakam James,G., Koornneef,M., Ossowski,S. *et al.* (2016) Chromosome-level assembly of Arabidopsis thaliana Ler reveals the extent of translocation and inversion polymorphisms. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E4052–E4060.
56. Stanley,E.C., Azzinaro,P.A., Vierra,D.A., Howlett,N.G. and Irvine,S.Q. (2016) The Simple Chordate Ciona intestinalis Has a Reduced Complement of Genes Associated with Fanconi Anemia. *Evol. Bioinform. Online*, **12**, 133–148.
57. Wang,S., Wang,R., Peralta,C., Yaseen,A. and Pavletich,N.P. (2021) Structure of the FA core ubiquitin ligase closing the ID clamp on DNA. *Nat. Struct. Mol. Biol.*, **28**, 300–309.
58. Cheng,C.Y., Krishnakumar,V., Chan,A.P., Thibaud-Nissen,F., Schobel,S. and Town,C.D. (2017) Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J.*, **89**, 789–804.
59. Klepikova,A.V., Kasianov,A.S., Gerasimov,E.S., Logacheva,M.D. and Penin,A.A. (2016) A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *Plant J.*, **88**, 1058–1070.
60. Zhang,X.Y., Langenick,J., Traynor,D., Babu,M.M., Kay,R.R. and Patel,K.J. (2009) Xpf and not the Fanconi anaemia proteins or Rev3 accounts for the extreme resistance to cisplatin in Dictyostelium discoideum. *PLoS Genet.*, **5**, e1000645.
61. Burki,F., Roger,A.J., Brown,M.W. and Simpson,A.G.B. (2020) The new tree of eukaryotes. *Trends Ecol. Evol.*, **35**, 43–55.
62. Kurzbauer,M.T., Pradillo,M., Kerzendorfer,C., Sims,J., Ladurner,R., Oliver,C., Janisiw,M.P., Mosiolek,M., Schweizer,D., Copenhaver,G.P. *et al.* (2018) *Arabidopsis thaliana* FANCD2 promotes meiotic crossover formation. *Plant Cell*, **30**, 415–428.
63. Walden,H. and Deans,A.J. (2014) The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. *Annu. Rev. Biophys.*, **43**, 257–278.
64. Ito,S. and Nishino,T. (2021) Structural analysis of the chicken FANCM-MHF complex and its stability. *Acta Crystallogr. F Struct. Biol. Commun.*, **77**, 1–7.