Arabidopsis HAP2/GCS1 is a gamete fusion protein homologous to somatic and viral fusogens

Clari Valansi,[1]* David Moi,[2]* Evgenia Leikina,[3] Elena Matveev,[1] Martín Graña,[4] Leonid V. Chernomordik,[3] Héctor Romero,[5] Pablo S. Aguilar,[2] and Benjamin Podbilewicz[1]

[1]Department of Biology, Technion- Israel Institute of Technology, Haifa 32000, Israel
[2]Laboratorio de Biología Celular de Membranas, Instituto de Investigaciones Biotecnologicas "Dr. Rodolfo A. Ugalde," Universidad Nacional de San Martin, Buenos Aires, CP1650, Argentina
[3]Section on Membrane Biology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892
[4]Unidad de Bioinformática, Institut Pasteur Montevideo, Montevideo 11400, Uruguay
[5]Laboratorio de Organización y Evolución del Genoma, Unidad de Genómica Evolutiva, Departamento Ecología y Evolución, Facultad de Ciencias/Centro Universitario Regional del Este, Universidad de la República, Montevideo 11400, Uruguay

Cell–cell fusion is inherent to sexual reproduction. Loss of HAPLESS 2/GENERATIVE CELL SPECIFIC 1 (HAP2/GCS1) proteins results in gamete fusion failure in diverse organisms, but their exact role is unclear. In this study, we show that *Arabidopsis thaliana* HAP2/GCS1 is sufficient to promote mammalian cell–cell fusion. Hemifusion and complete fusion depend on HAP2/GCS1 presence in both fusing cells. Furthermore, expression of HAP2 on the surface of pseudotyped vesicular stomatitis virus results in homotypic virus–cell fusion. We demonstrate that the *Caenorhabditis elegans* Epithelial Fusion Failure 1 (EFF-1) somatic cell fusogen can replace HAP2/GCS1 in one of the fusing membranes, indicating that HAP2/GCS1 and EFF-1 share a similar fusion mechanism. Structural modeling of the HAP2/GCS1 protein family predicts that they are homologous to EFF-1 and viral class II fusion proteins (e.g., Zika virus). We name this superfamily Fusexins: fusion proteins essential for sexual reproduction and exoplasmic merger of plasma membranes. We suggest a common origin and evolution of sexual reproduction, enveloped virus entry into cells, and somatic cell fusion.

Introduction

Although proteins mediating cell–cell fusion in tissues have been demonstrated in the placenta of mammals (Syncytins) and in organs of invertebrates (e.g., Epithelial Fusion Failure 1 [EFF-1] in *Caenorhabditis elegans*), the machinery mediating sperm–egg fusion remains unknown (Aguilar et al., 2013; Bianchi et al., 2014; Maruyama et al., 2016). HAPLESS 2/GENERATIVE CELL SPECIFIC 1 (HAP2/GCS1) proteins have been implicated as potential gamete fusogens in *Arabidopsis thaliana* (Johnson et al., 2004; Mori et al., 2006; von Besser et al., 2006), *Chlamydomonas reinhardtii* (Liu et al., 2008), *Tetrahymena thermophila* (Cole et al., 2014), *Dictyostelium discoideum* (Okamoto et al., 2016), and *Plasmodium falciparum* (Liu et al., 2008). However, the precise function of HAP2/GCS1 in gamete fusion is unknown. So far, there is no functional or structural evidence indicating HAP2/GCS1 is directly involved in cell–cell fusion. Proteins may function as direct fusogens, or alternatively, they may affect communication or intimate adhesion before fusion takes place, as demonstrated for other gamete fusion candidates such as Juno and Izumo receptors (Bianchi et al., 2014).

Results and discussion

To determine whether HAP2/GCS1 is an authentic fusion protein, we first tested whether *A. thaliana* HAP2 (AtHAP2) could fuse heterologous cells that normally do not fuse. For this, we transfected BHK cells with plasmids encoding AtHAP2, EFF-1, or RFP or GFP as negative controls and assayed the extent of cell–cell fusion (Fig. 1 A). In controls, when BHK cells were transfected with cytoplasmic RFP (RFPcyto-BHK) and mixed with GFP-transfected BHK cells (GFP-BHK; Fig. 1 B, i), ~5% of cells (red or green, respectively) had two nuclei because of cell division, and only 1.5% of the cells expressed both GFP and RFPcyto out of the total GFP/RFPcyto-expressing cells in contact (Fig. 1 C). This apparent cytoplasmic content mixing could be because of phagocytosis of fluorescent apoptotic bodies or background fusion. In contrast, when AtHAP2 was transfected into BHK cells with either RFPcyto or GFP and the transfected cells were coincubated, we observed a mean multinucleation of 33 ± 3 and $41.3 \pm 1.3\%$ (green or red) and cytoplasmic content mixing in $11.3 \pm 0.9\%$ in three independent experiments (Fig. 1, B [ii and iv] and C). Similar results were obtained using the previously defined *C. elegans*' somatic cell fusogen EFF-1 (Fig. 1, B [iii] and C). To test whether some
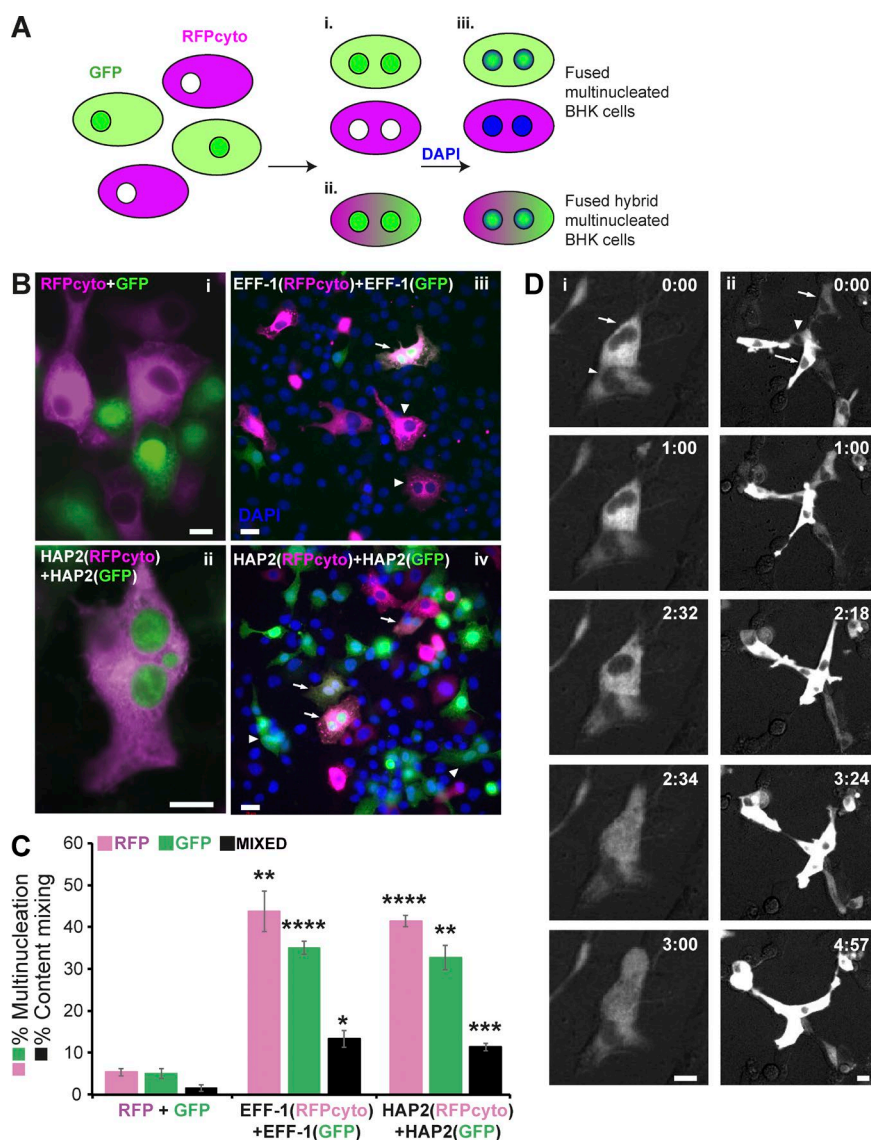
Figure 1. *A. thaliana* HAP2 is sufficient to fuse mammalian BHK cells. (A) BHK cell–cell fusion assay: after discarding a possible failure in cell division (Table S1), cell–cell fusion is measured by the appearance of multinucleated cells labeled with either RFPcyto (magenta) or nuclear and cytoplasmic GFP (green; i). (ii) Fusion is also indicated by the appearance of multinucleated cells containing nuclear GFP and fluorescence from both RFPcyto and GFP in the cytoplasm. (iii) Nuclei are labeled with DAPI (blue) after fixation and permeabilization of the cells. (B, i) RFPcyto + GFP: negative control shows mononucleated cells expressing RFPcyto (magenta) or nuclear and cytoplasmic GFP (green). (ii) HAP2(RFPcyto) + HAP2(GFP): BHK cells were transfected with AtHAP2 and GFP (green) or RFPcyto (magenta); merged image of hybrid cell that contains mixed cytoplasm and three nuclei. (iii) EFF-1(RFPcyto) + EFF-1(GFP): hybrid binucleate cell emerged after EFF-1 expression and mixing of magenta and green cells (arrow). EFF-1(RFPcyto) binucleate cells (arrowheads). (iv) HAP2(RFPcyto) + HAP2(GFP): heterokaryons (hybrids) express magenta cytoplasm and green nuclei and cytoplasm (arrows). Multinucleate green cells (arrowheads). Bars: (B, i and ii) 10 µm; (B, iii and iv) 20 µm. (C) Quantification of multinucleation and content-mixing experiments. Magenta and green bars represent the fraction of multinucleated cells (two nuclei or higher) out of all the cells in contact (magenta or green, respectively). Black bars represent the RFPcyto and GFP content-mixing index. The fusion and mixing indexes are presented as means ± SEM of three independent experiments. Total number of nuclei counted in multinucleated cells and in cells in contact $n \geq 1,000$ for each experimental condition. Used unpaired *t* test comparing each color (RFPcyto, GFP, or mixed) for EFF-1 and HAP2 to the negative control (RFPcyto+GFP). *, $P < 0.01$; **, $P < 0.005$; ***, $P < 0.001$; ****, $P < 0.0005$. (D) Still images from time-lapse experiments reveal merging of two mononucleated (i) and three cells (ii) expressing RFPcyto and HAP2 (arrows and arrowheads, respectively). Time indicated in hours :minutes (see Videos 1 and 2 for panels i and ii in D, respectively). Note that the top nucleus (arrow in D, i) disappears because of defocus at 2:34. Two nuclei are out of focus at 4:57 (D, ii, bottom; see Fig. S1 A). Bars, 20 µm.

of the multinucleated cells resulted from faulty cell division (e.g., nuclear division without cytokinesis), we incubated BHK-HAP2 cells with 5-fluoro-2'-deoxyuridine (FdUrd) to arrest the cell cycle at the G$_1$/S transition. We found that FdUrd failed to decrease multinucleation in HAP2-BHK cells and actually doubled the multinucleation, probably because of a block in cell division of multinucleated cells formed by cell–cell fusion. In contrast, FdUrd did not significantly affect nuclei content of BHK cells that were transfected with the GFP plasmid only (Table S1). Time-lapse microscopy experiments showed that mononucleated BHK-HAP2(RFPcyto) cells formed syncytia by merging their cytoplasms (Fig. 1 D; Fig. S1 A; and Videos 1, 2, 3, and 4). Thus, AtHAP2 expression in BHK cells is sufficient to promote cell–cell fusion, defining this protein as a bona fide fusogen.

We asked whether HAP2/GCS1 family members display any similarity to known fusogenic proteins. HAP2/GCS1 are type I membrane glycoproteins composed of an N-terminal signal peptide, a large ectodomain, and a C-terminal cytoplasmic tail (Fig. 2 A). The ectodomain contains a conserved 50-aa region (Pfam PF10699, H/G domain; Fig. 2 A) that is unique to the HAP2/GCS1 protein family (Wong and Johnson, 2010). In search of structural similarities, we filtered and compiled ectodomains from all members of the HAP2/GCS1 family and subjected this dataset to two homology detection and structure prediction algorithms, HHblits (Söding, 2005; Remmert et al., 2012), and LOMETS (Wu and Zhang, 2007). We detected structural similarity to ectodomains of the EFF-1 protein from *C. elegans* (Pérez-Vargas et al., 2014) and class II viral fusogenic glycoproteins (Harrison, 2008; White et al., 2008; Kielian et al., 2010; Igonet and Rey, 2012; Podbilewicz, 2014) from three viral families: Flaviviridae, Togaviridae, and Bunyaviridae (Fig. S2, A and B). Both methods detected similarity to class II fusogens throughout the entire HAP2/GCS1 family, consistent with their expected role in gamete fusion in evolutionary distant species (Fig. S2, C and D). To generate structural models of AtHAP2, we used I-TASSER (Yang et al., 2015). The overall fold of the AtHAP2 ectodomain model showed the canonical architecture of class II fusogens with three domains: a β-barrel domain I (DI), a mostly β-stranded
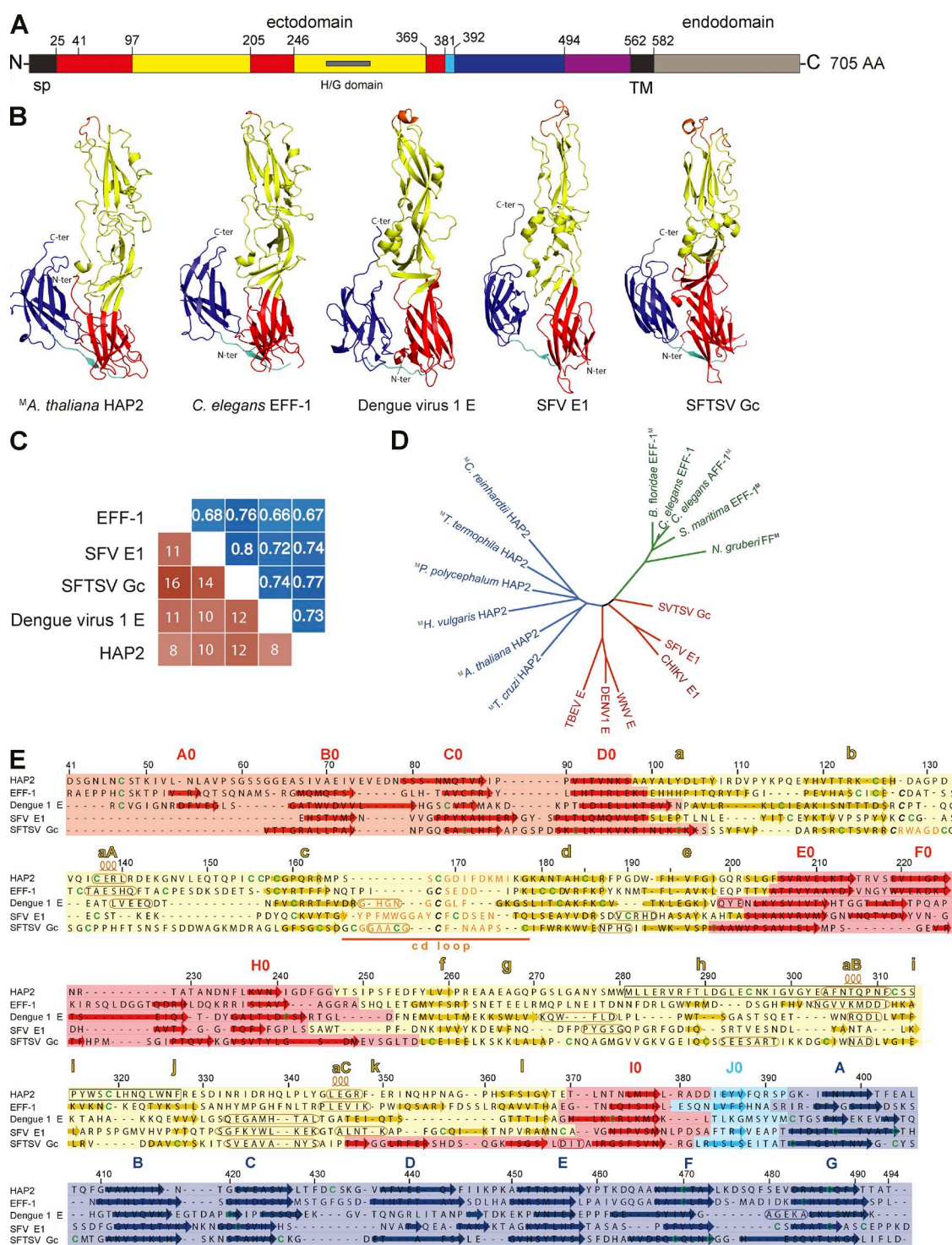
Figure 2.   **Model of the 3D structure of *A. thaliana* HAP2 ectodomain indicates it is a class II fusogen.** (A) Diagram of *A. thaliana* HAP2 protein colored by domains according to the ectodomain modeled structure: signal peptide (sp) and transmembrane domain (TM), black; domain I, red; domain II, yellow; domain III, blue; domain I–III linker, cyan; stem, magenta; and intracellular domain, gray. (B) Cartoon of AtHAP2 ectodomain modeled structure (residues 41–494), alongside with *C. elegans* EFF-1 (PDB 4OJC), dengue virus E glycoprotein (PDB 4GSX), Semliki Forest virus E1 glycoprotein (SFV E1; PDB 1RER), and severe fever with thrombocytopenia syndrome virus glycoprotein Gc (SFTSV Gc; PDB 5G47) class II fusion proteins. Structures are colored according to domains as in A and E; cd loops are shown in orange (facing up). (C) Structural similarity scores between experimental and modeled fusogens computed after flexible structural alignment. Blue matrix: TM-scores (values >0.5 are indicative of significant structural similarity); orange matrix: z-scores (values >2 are considered as indicative of significant structural similarity). (D) Unrooted tree inferred using a distance matrix. Superscript gray M indicates modeled structure; colors are HAP2, blue; EFF-1/AFF-1/FF, green; class II viral fusogens, red (Fig. S3 A). (E) Structure-based alignment of AtHAP2 ectodomain with fusion proteins shown in B. Background colors indicate the domains organization (as in A and B); arrows in color and rounded rectangles denote β sheets and α helices, respectively. The black box in AtHAP2 sequence marks the HAP2/GCS1 domain; the cd loops are marked in orange. Cysteines involved in conserved disulfide bonds stabilizing the cd loop are denoted in bold italics.
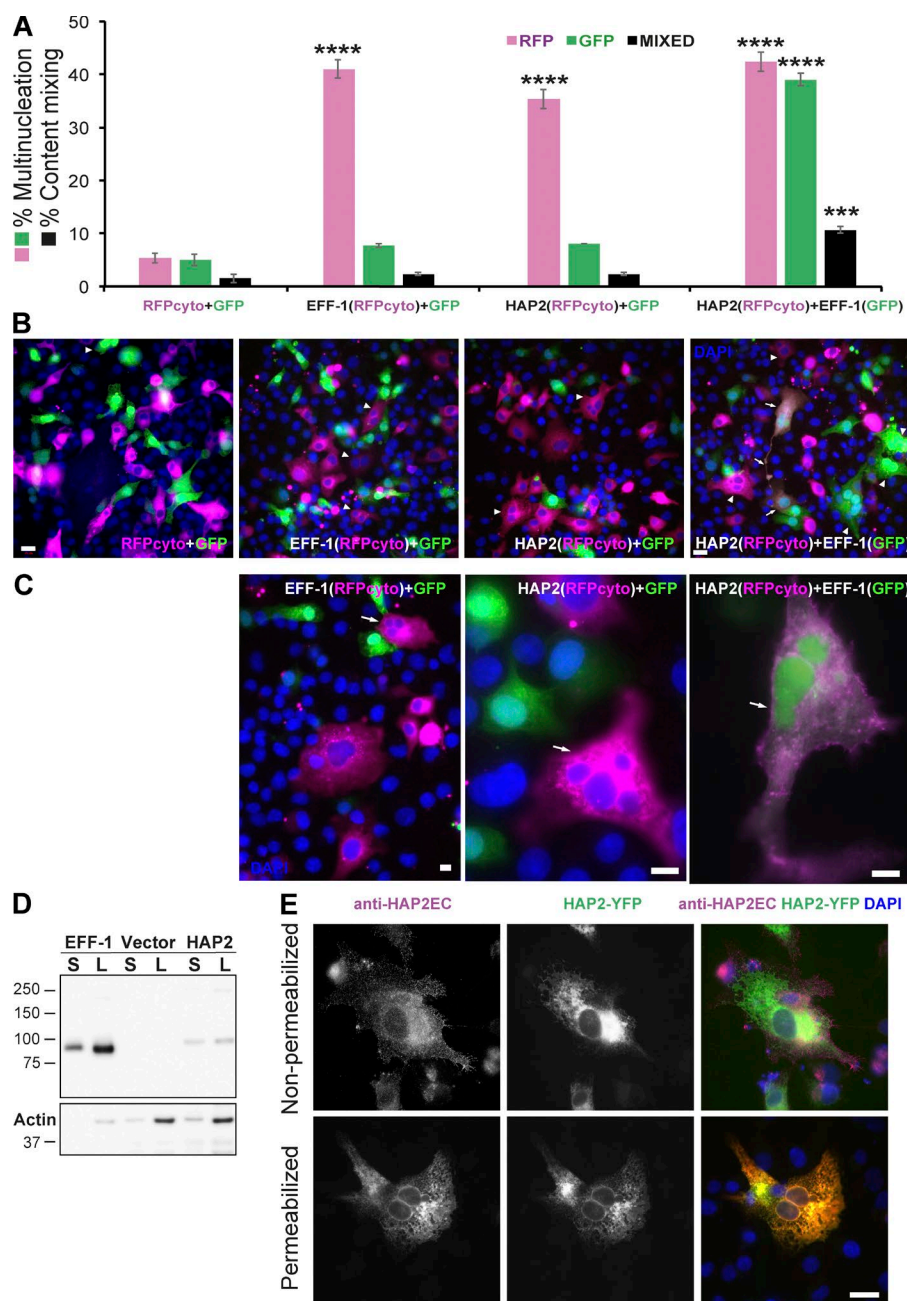
Figure 3. HAP2 is a bilateral fusogen. (A) Quantification of content mixing experiments showing that HAP2 and EFF-1 are required in both interacting cells to form hybrids. Control RFPcyto + GFP is the same as in Fig. 1 B. The fusion and mixing indexes are presented as means ± SEM of three independent experiments. Total number of nuclei counted in multinucleated cells and in cells in contact, $n \geq$ 1,000 for each experimental condition. ***, $P < 0.001$; ****, $P < 0.0005$. $P > 0.05$ is not significant. (B) Representative fields used to determine percentages of multinucleation and content mixing. RFPcyto + GFP: mixed control cells; nuclear staining DAPI (blue). Dividing green cell with two nuclei is marked by arrowhead. EFF-1(RFPcyto) + GFP: BHK–EFF-1 (magenta) do not mix with GFP-expressing cells, revealing that EFF-1–mediated fusion is bilateral (homotypic). Multinucleate cells expressing RFPcyto (arrowheads). HAP2(RFPcyto) + GFP: BHK-HAP2 multinucleation (arrowheads) and failure to mix with GFP-expressing cells revealing HAP2-mediated fusion is bilateral. HAP2(RFPcyto) + EFF-1(GFP): hybrids (arrows) between BHK–EFF-1 (green) and BHK-HAP2 (magenta) reveal heterotypic merger of cells. Multinucleated green or red cells (arrowheads). Some hybrids are mononucleate probably because of cell division or nuclear fusion after merger. Bars, 20 µm. (C) Examples of multinucleate cells containing two to four nuclei are marked with arrows. The cells are able to divide after fusion; therefore, the number of nuclei per cells is usually smaller than six. EFF-1(RFPcyto) + GFP and HAP2(RFPcyto) + GFP images show no mixing indicating that HAP2-mediated fusion is bilateral (homotypic) in BHK cells. Bars, 10 µm. (D) Immunoblot of vector (negative control) and EFF-1 and HAP2 proteins carrying a V5 epitope fused to the cytoplasmic tail. Surface biotinylation of BHK cells expressing the different proteins. S indicates surface expression after affinity purification using neutravidin agarose beads; L is lysate. The amount of sample of HAP2 "S" is 600 times higher than EFF-1. The amount of lysate for HAP2 "L" is 12 times higher than for EFF-1. (E) Immunofluorescence of HAP2-YFP shows surface expression in cells revealed by anti-HAP2Extracellular (HAP2EC) polyclonal antibody. Without permeabilization, 20% of the cells expressing HAP2-YFP showed punctate expression on the surface; merged image (top, magenta). Untransfected BHK cells showed no immunoreactivity. After permeabilization with detergent, there is colocalization between immunostaining with anti-HAP2EC and the YFP signal revealing the specificity of the antibody that recognizes HAP2 in the reticular ER and perinuclear localization (merged image; bottom, yellow). Immunofluorescence in permeabilized BHK-HAP2-V5 cells using anti-V5 antibody shows similar localization (Fig. S1, B and C). Bar, 20 µm.

elongated domain (DII), and the Ig-like C2-set topology module (DIII), separated by a linker from DI (Fig. 2 B). Global protein structure comparisons among these proteins, as measured by Z-score (Holm et al., 2008) and TM-score (Wu et al., 2007), indicate high similarity, with scores typical for proteins belonging to the same fold (Fig. 2 C and Fig. S3 A). Structural models with similar architecture were also obtained for HAP2/GCS1 proteins from several species (Fig. S3 C). An unrooted tree inferred from a structural dissimilarity matrix shows the structural relationship between families (Fig. 2 D). Although amino acid sequence conservation among AtHAP2, somatic, and viral class II fusogens is low, most β-strands of AtHAP2

ectodomain are conserved and arranged in the same way as in *C. elegans* EFF-1 and viral fusogens (Fig. 2 E; Pérez-Vargas et al., 2014). AtHAP2 also shares conserved cysteine residues with class II fusogens and a prominent loop at the tip of DII (the cd loop; Fig. 2 E). In viral fusogens, this loop is highly hydrophobic and necessary for anchoring at the host cell membrane (Podbilewicz, 2014), whereas in EFF-1, it is negatively charged (Pérez-Vargas et al., 2014). Unlike EFF-1, the AtHAP2 predicted cd loop is hydrophobic and flanked by charged residues (Fig. S3 D). Two additional loops neighbor the AtHAP2 cd loop (Fig. 2). One of these loops (between the i and j β-strands) forms part of the conserved H/G domain,

is hydrophobic, and point mutations at equivalent residues of *C. reinhardtii* HAP2/GCS1 block trafficking to the plasma membrane (Liu et al., 2015).

Class II fusion proteins mediate exoplasmic membrane fusion either unilaterally (e.g., from the virus envelope to the cell membrane) or bilaterally (e.g., EFF-1–mediated cell fusion; Podbilewicz et al., 2006; Podbilewicz, 2014). HAP2 is essential in sperm for double fertilization in *A. thaliana* (Johnson et al., 2004; Mori et al., 2006; von Besser et al., 2006). In *C. reinhardtii* and *P. falciparum*, HAP2 is also active in male gametes (Liu et al., 2008). However, in the seven-sexed *T. thermophila*, efficient fertilization requires the presence of HAP2 in both fusing gametes (Cole et al., 2014). To test whether HAP2-mediated cell fusion requires HAP2 presence in either one or both fusing cells, we mixed HAP2-BHK-RFPcyto cells with BHK-GFP cells and determined the percentage of multinucleation in cells expressing RFPcyto, GFP, or both. Whereas HAP2-BHK-RFPcyto cells had 35 ± 1.9% multinucleation, cytoplasmic content mixing between HAP2-BHK-RFPcyto and BHK-GFP cells was only 2 ± 0.3%, undistinguishable from the negative control transfections (Fig. 3 A). These findings indicate that AtHAP2 is required in both cells for fusion to occur (Fig. 3 B).

We previously demonstrated that EFF-1 and its paralog Anchor cell Fusion Failure 1 (AFF-1) can fuse BHK cells in a heterotypic way (Avinoam et al., 2011). To test whether AtHAP2 can promote cell fusion in trans with the nematode protein EFF-1, we mixed HAP2-BHK-RFPcyto cells with EFF-1–BHK-GFP cells and looked for cells containing both RFPcyto and GFP. We found that hybrids formed as efficiently as with EFF-1 in all fusing cells (10.7 ± 0.7%; Fig. 3, A–C). Thus, EFF-1 can substitute for HAP2 in one of the cell membranes.

The extent of cell fusion is proportional to the level of surface expression of fusion proteins (e.g., EFF-1 and AFF-1; Podbilewicz et al., 2006; Gattegno et al., 2007; Sapir et al., 2007; Avinoam et al., 2011). To determine whether AtHAP2 is expressed on the plasma membrane, we used surface biotinylation followed by immunoblotting and immunofluorescence using an antibody against the extracellular domain of HAP2 in intact and permeabilized BHK cells. After surface biotinylation, we detected low amounts of HAP2-V5 on the surface compared with EFF-1–V5 (Fig. 3 D). Immunofluorescence showed that AtHAP2 localized to puncta on the surface of 20% of the expressing BHK-AtHAP2-YFP cells (Fig. 3 E and Fig. S1, B and C). Because lower surface expression of AtHAP2, compared with EFF-1, resulted in similar fusogenic activity of both proteins, we conclude that AtHAP2 is a more efficient fusion protein compared with EFF-1.

Previously, we demonstrated that eukaryotic fusogens can be assayed efficiently by replacing the vesicular stomatitis virus fusion glycoprotein G (VSVG) with a foreign fusogen and testing for the infectivity of the resulting pseudovirus (Avinoam et al., 2011). To test whether AtHAP2 can functionally substitute for VSVG, the fusogen glycoprotein G gene was deleted (VSVΔG) and replaced by HAP2 (VSVΔG-HAP2; Fig. 4 A). We first infected BHK-HAP2 cells with VSVG-complemented VSVΔG recombinant virus (VSVΔG-G) pseudoviruses and collected VSVΔG-HAP2 pseudoviruses from the supernatants. As controls, we generated VSVΔG–EFF-1 particles (Avinoam et al., 2011). To test whether VSVΔG-HAP2 can infect cells, we inoculated BHK cells, naive or expressing HAP2, or EFF-1 (Fig. 4 B).
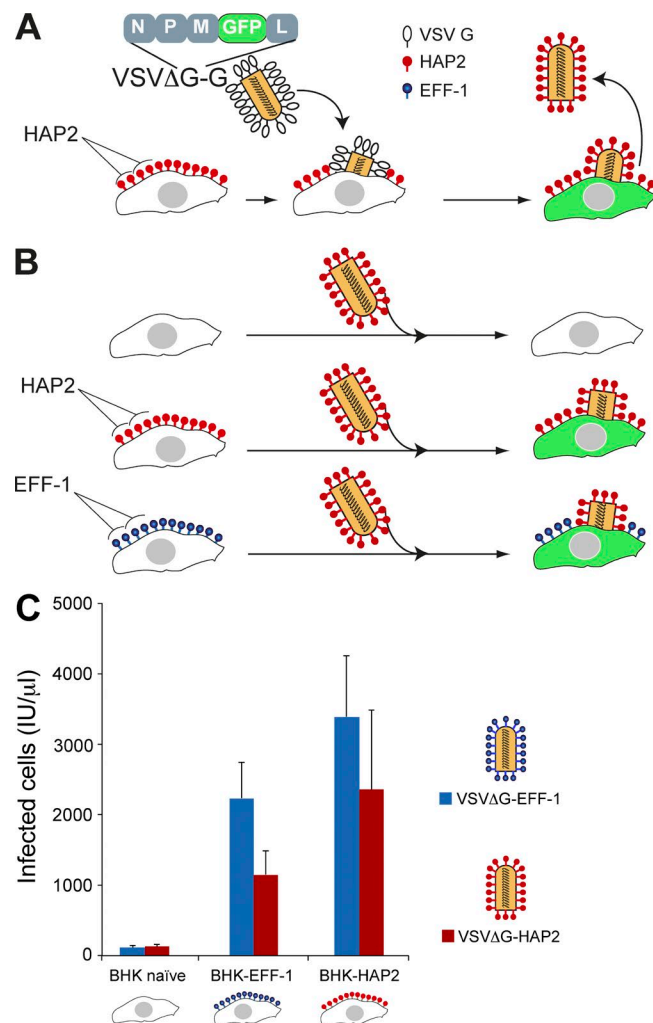
Figure 4. **HAP2 can fuse the VSVΔG pseudovirus to target cells.** (A) Cartoon illustrates the generation of VSVΔG-HAP2 pseudoviruses. Transfected BHK cells express HAP2 protein on the surface and were infected with VSVΔG-G. The viral genome encodes GFP, replacing VSVG. Infection results in viral-induced expression of GFP by target cells (green cytoplasm). VSVΔG-HAP2 pseudoviruses were collected from the supernatant. (B) The activity of VSVΔG-HAP2 was tested on BHK (untransfected, naive), BHK-HAP2, and BHK-EFF-1 cells. Figure adapted with permission from Avinoam et al. (2011). (C) Titers of VSVΔG pseudoviruses. The type of protein on the viral membrane (EFF-1 or HAP2) and on the BHK target cells (naive, EFF-1, or HAP2) is indicated. Titers in infectious units (IU) represent the number of cells expressing GFP per microliter 24 h after virus inoculation. Data are mean ± SEM (n = 3 experiments). We found no significant difference between infection with VSVΔG-HAP2 and VSVΔG–EFF-1 of the different BHK target cells (two-tailed paired *t* test).

We observed a 7- to 70-fold increase in infection efficiency in the fusogen-expressing cells compared with naive cells (Fig. 4 C). Finally, to determine whether EFF-1 and HAP2 can interact through a bilateral mechanism in this virus–cell fusion system, we infected BHK–EFF-1 with VSVΔG-HAP2 and obtained similar infection efficiency to homotypic infection of BHK-HAP2. Likewise, VSVΔG–EFF-1 pseudotyped viruses showed similar infection efficiencies when infecting BHK-HAP2 or BHK–EFF-1 cells (Fig. 4 C). These finding reinforce our conclusion from cell–cell fusion assays (Fig. 3) and further suggest that HAP2 and EFF-1 directly interact to mediate heterotypic membrane fusion.
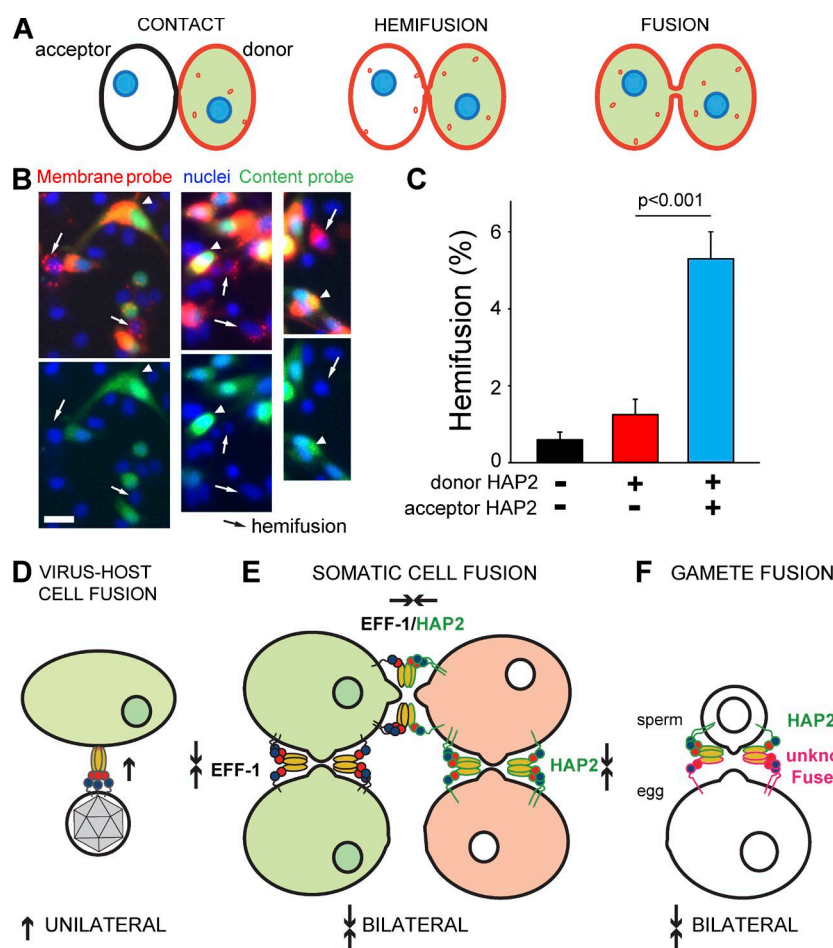
Figure 5. **Hemifusion in HAP2-mediated fusion is bilateral, and Fusexins use divergent mechanisms.** (A) Cartoon illustrates the hemifusion assay in which one cell type, labeled with two fluorescent probes, acts as the "donor" cell and the unlabeled cell acts as "acceptor." Because of internalization of fluorescent lipid from the plasma membrane, by the time we score fusion, this probe mostly labels intracellular membranes. (B) Fluorescence microscopy images of AtHAP2-transfected cells labeled with both cell tracker (green, content probe) and Vybrant DiI (red, membrane probe) coplated with unlabeled AtHAP2-transfected cells. (top) Green cell tracker, DiI (red), and Hoechst 33342 (blue). (bottom) Green cell tracker and Hoechst 33342 (blue). Hemifusion event is detected as an appearance of a cell (marked by arrows) that acquired only membrane probe apparently from an adjacent double-labeled cell (arrowheads). Bar, 20 µm. (C) Hemifusion extents quantified as the ratio between numbers of cells labeled only with membrane probe and numbers of cells labeled with both membrane and content probes. The results are means ± SEM ($n = 3$). (D) Viral class II trimeric fusion proteins (viral Fusexins) have a unilateral fusion mechanism, and the Fusexin is present only in the virus' envelope or in one cell during cell–cell fusion. (E) Somatic cellular Fusexins (e.g., EFF-1 and AFF-1; on green cells) use a bilateral homotypic mechanism, and the model proposes Fusexin has to be in both cells, forming trans-trimeric complexes (Podbilewicz et al., 2006; Pérez-Vargas et al., 2014). Our results suggest a mechanistic model in which sperm AtHAP2 fuses animal cells using a homotypic design (bilateral; pink skin). HAP2 and EFF-1 can fuse cells in trans using a heterotypic mechanism. (F) We hypothesize that the egg of *A. thaliana* (the central cell too) expresses an unidentified Fusexin that interacts with sperm HAP2.

We then examined whether AtHAP2-mediated fusion proceeds via hemifusion, a fusion intermediate in which outer leaflets of two membranes are already merged but the inner leaflets remain distinct (Chernomordik and Kozlov, 2005). To this end, cells transfected with AtHAP2 were labeled with both Cell Tracker as a cytosolic probe and DiI as a membrane probe and plated together with transfected but unlabeled cells. In this experimental design, cell fusion caught at a stage between hemifusion and fusion is expected to produce cells that acquired membrane, but not content, probe (Fig. 5 A). Indeed, we observed the appearance of such cells (Fig. 5 B, arrows). Importantly, the number of cells labeled with only membrane probe was much lower in experiments in which the labeled cells expressing AtHAP2 were coplated with unlabeled and nontransfected cells (Fig. 5 C). The apparent hemifusion in this control (~1.25%) could reflect background fusion or overlapping cells. Our results indicate that HAP2-mediated fusion proceeds through hemifusion intermediates. Our data are also the first evidence that a homotypic fusogen has to be present in both membranes to mediate even hemifusion, a fusion stage that is less energy demanding than opening and expansion of a fusion pore (Chernomordik and Kozlov, 2005).

Together, our findings indicate a somatic cell-like bilateral mechanism for AtHAP2 rather than a unilateral viral-like mechanism. Such bilateral mechanism is likely to operate not only in flowering plants but also in other organisms expressing members of the HAP2/GCS1 family (Fig. 5, D–F).

The observed structural and functional similarities among the gamete HAP2/GCS1 proteins, EFF-1 fusogen, and class II viral fusion proteins could result from either common ancestry or convergence. Our data suggest a common ancestry based on the following observations: first, HAP2, EFF-1, and class II viral fusion proteins share a common domain architecture (Fig. 2). Second, proteins of the same fold are likely to be homologues, with only a small fraction having different ancestries (Forslund et al., 2008). Third, comparisons of hidden Markov model (HMM)–based profiles of the complete ectodomain from different members of class II viral fusogens, HAP2, and EFF-1/AFF-1 (FF) show low but significant amino acid sequence similarity among members of different groups (Fig. S3 B). Altogether, the combined conservation of structure, sequence, and function makes a robust case for considering these proteins as diverged from a common ancestor.

Based on the ancient origin and function of exoplasmic membrane fusion, we suggest naming this superfamily of fusion proteins the Fusexins (fusion proteins essential for sexual reproduction and exoplasmic merger of plasma membranes). In conclusion, we have demonstrated the fusogenic activity of a fertilization protein and found a molecular link unifying sexual, somatic, and viral membrane fusion. We hypothesize that Fusexins are ancient proteins, and their emergence, either in viruses or cells, may have been a key innovation spurring sexual reproduction, exoplasmic somatic membrane fusion, and the rise of eukaryotes >2.5 billion years ago.

Materials and methods

Homology detection

To detect structural homology in the HAP2/GCS1 sequences annotated in the InterPro database were taken as the HAP2 dataset (IPR018928). A total of 299 aa sequences were analyzed with TOPCONS (Tsirigos et al., 2015) to predict membrane topology. The resulting 299-ectodomain amino acid sequences were then used as queries against HHblits (Remmert et al., 2012) and LOMETS (Wu and Zhang, 2007). The Uniprot20 database was used to build HMMs with one iteration of HHblits using default parameters. A subsequent search was performed using these HMMs against the Protein Data Bank (PDB) 70 database with default parameters to generate the final structural prediction using HMM–HMM alignment. To collect the results of the HHblits pipeline on all sequences, the output files were parsed with a script using CSB toolbox (Kalev et al., 2012). LOMETS, a meta algorithm that is part of the I-TASSER suite (Yang et al., 2015), was also used to predict structural homology for the 299 HAP2/GCS1 ectodomains. The standalone version uses six structural threading algorithms that are all variations of MUSTER (Wu and Zhang, 2008). LOMETS was run with default parameters for each HAP2 ectodomain sequence. LOMETS output files were parsed with a custom parser written in python to compile alignment template codes, columns aligned, and z-scores (supplemental data files 2 and 3). All structural predictions using LOMETS and HHblits were performed locally. TOPCONS topology predictions were run on the TOPCONS server. Complete lists of structural prediction results are available in Table S2.

Models generation

Structural models were generated for a taxonomically diverse subset of ectodomain sequences using the I-TASSER server (Yang et al., 2015). EFF-1 and HAP2 ectodomains were prepared by removing signal sequences, transmembrane domains, and C-terminal intracellular domains in the same way as for the structural prediction pipelines. Models were built from the ectodomains of sequences from the HAP2/GCS1 family available in the UniProt database for *A. thaliana* (F4JP36), *Hydra vulgaris* (A3FEQ2), *Physarium polycephalum* (Q2PGG5), *C. reinhardtii* (A4GRC6), *Trypanosoma brucei gambiense* (D0A4L4), and *T. thermophila* (A0A060A682). Additional models for the AFF-1/EFF-1 family were built from *Naegleria gruberi* (D2W008), *Strigamia maritima* (T1ISJ2), *C. elegans* (G5EGL9), and *Branchiostoma floridae* (C3Y7J9) ectodomains. Structural constraints based on probable disulfide bonds were used as additional inputs on the I-TASSER server for subsequent models for the *A. thaliana* HAP2 ectodomain. These bonds were inferred from residue conservation in MSAs of HAP2/GCS1 ectodomains and visual inspection of the preliminary model. I-TASSER was run a second time on the *A. thaliana* ectodomain sequence using these structural constraints to generate the final model presented in this publication. All of the PDB models generated using I-TASSER are available in the supplemental data file 1.

Structure-based multiple sequence alignment

Structural alignments were used to generate high quality MSAs and pairwise superpositions of structures for use with structural similarity metrics. FATCAT (Ye and Godzik, 2004) pairwise alignments were run on all pairs of monomers in the I-TASSER models and crystal structures. Using a custom parser of the FATCAT output, pairwise alignments were generated in Fasta format. These alignments were iteratively merged using Clustal omega (Sievers et al., 2011) so that in each merger, at least one sequence was shared between the Fasta files being merged until a global alignment containing at least one of each of the sequences for the structures being aligned is produced. These final Fasta format alignments contain multiple sequence alignment possibilities for the set structures

fed to the algorithm. Because all of the possibilities of sequence alignment are relatively similar, the final version of the alignment is chosen in Jalview (Waterhouse et al., 2009) by manual inspection.

Structural similarity and structural phylogeny

Structural similarity matrices were built using the Dali Z-score (Holm and Rosenström, 2010) and the TM-score (Zhang and Skolnick, 2005) using pairwise superpositions generated by FATCAT. It has been shown that Z-scores >2 (Holm and Rosenström, 2010) and TM-scores >0.5 (Xu and Zhang, 2010) are indicative of proteins belonging to the same folds, usually homologous. TM-score matrix was transformed into a distance matrix with the metric distance = $1 -$ TM-score. This structure distance matrix was used to infer a phylogeny of ectodomains using FASTME (Lefort et al., 2015) on the default settings. The automation script made for the generation of structure-based phylogenies is available in the supplemental data file 2.

Profile-based phylogeny

A taxonomically diverse set of ectodomain sequences of fusogens were used as search queries for HHblits against the Uniprot20 database to generate profiles, as HMMs incorporating also secondary structure information. For model construction realignment, greediness (mact) was set to 10^{-1}, and HHblits was run for three iterations on otherwise default parameters. These models were used in an all-versus-all search using HHsearch. Realignment greediness (mact) and secondary structure weight (ssw) in realignment in the all-versus-all HHsearch alignments were set to 10^{-4} and 3×10^{-1}, respectively. The automation script made for the generation of profile-based phylogenies is available in the supplemental data file 2.

Installation and configuration of structural similarity analysis and profile-to-profile tools

Python scripts used for structural and profile-to-profile analyses, are available as supplemental text files. They depend on SciPy, NumPy, Pandas, Biopython, and the CSB toolbox python libraries. The structural phylogeny script requires the installation of TMalign and the PDB protein comparison tool (http://www.rcsb.org/pdb/workbench/workbench.do?action=menu). The profile-to-profile analysis script requires the installation of HHblits and HHsearch along with the Uniprot20 and PDB70 databases. These scripts were built for use in an Ubuntu 14.04LTS environment and have not been tested using other operating systems.

Electrostatic calculations

To calculate electrostatic properties of the protein surfaces, the Adaptive Poisson-Boltzmann Solver (Baker et al., 2001) was used with default parameters using atomic coordinates previously prepared with PDB2PQR (Dolinsky et al., 2007). Solvent-accessible surfaces were colored by electrostatic potential and finally rendered in PyMOL (https://sourceforge.net/projects/pymol/).

Cells and reagents

BHK cells were BHK-21 (ATCC). BHK cells were grown and maintained according to standard protocols using DMEM, supplemented with 10% FBS (Biological Industries), 100 U/ml penicillin, 100 µg/ml streptomycin (Biological Industries), 2 mM L-glutamine (Biological Industries), 1 mM sodium pyruvate (Gibco), and 30 mM Hepes buffer, pH 7.3, at 37°C (Biological Industries). Cells were grown at 37°C with 5% CO_2. Cells were transfected using a mixture of 8 µl Fugene HD (Promega) and 112 µl Opti-MEM (Gibco) with plasmids specified in the DNA constructs section in 1 ml for every 35-mm plate (Nunclon surface; Thermo Fisher Scientific) without replacing the medium (Avinoam et al., 2011).

DNA constructs

The *A. thaliana* HAP2 plasmids were derived from AtHAP2promoter::HAP2cds::YFP (PGL290; GenBank AF234315; von Besser et al., 2006) that fully rescues *hap2(−) A. thaliana* mutants (Wong et al., 2010). HAP2 ORF was amplified from PGL290 by PCR using primers 5′-GCACTAGTATGGTGAACGCGATTTTAAT-3′ and 5′-TCCCGCGGACTCTCACGTAGTCTTTGTT-3′. The amplified AtHAP2 fragment was subcloned into pIZT (Invitrogen) after SpeI and SacII digestion to obtain pIZT::AtHAP2-V5-HIS used for expression in insect cells. AtHAP2-V5 was amplified from pIZT::AtHAP2-V5-HIS using 5′-CGTTATCGATATGGTGAACGCGATTTTAAT-3′ and 5′-GCCCGGGTCACGTAGAATCGAGACCGA-3′. The AtHAP2-V5 fragment was subcloned into pCAGGS vector using ClaI and XmaI to obtain pCAGGS::AtHAP2-V5. To subclone AtHAP2-V5 into pGENE B–inducible system, pIZT::AtHAP2-V5 was PCR amplified using 5′-GAGGTACCATGGTGAACGCGATTTTAAT-3′ and 5′-CCACTAGTTCACGTAGAATCGAGACCGA-3′ primers and cloned into pGENE B between the KpnI and SpeI restriction sites to obtain pGENE B::AtHAP2-V5. AtHAP2-YFP was subcloned into pCAGGS between the restriction sites ClaI and XmaI. pGL290 was used as a template, and the insert was PCR amplified using primers 5′-CGTTATCGATATGGTGAACGCGATTTTAAT-3′ and 5′-GCCCGGGTTACTTGTACAGCTCGTCCA-3′. pGENE B::AtHAP2::YFP was obtained by PCR amplification of PGL290 using primers 5′-GAGGTACCATGGTGAACGCGATTTTAAT-3′ and 5′-CCACTAGTTCACGTAGAATCGAGACCGA-3′. The fragment was cloned into pGENE B after digestion with KpnI and SpeI. EFF-1A::V5 was PCR amplified from pIZT:EFF-1A::V5 (Podbilewicz et al., 2006) using primers OR54 (5′-TTAATTGGTACCACTATGGAACCGCCGTTTGAGTGG-3′) and OR55 (5′-AATTAAGCTAGCTCAACCGGTACGCGTAGAATCGAGACC-3′) and cloned into pGENE B using KpnI and SpeI.

For mifepristone-inducible expression in BHK cells, we used the GeneSwitch System (Wang et al., 1994) as recommended by Invitrogen (Avinoam et al., 2011). Cells were transfected with pGENE B::EFF-1A-V5, pGENE B::AtHAP2-V5, or pGENE B::AtHAP2::YFP. For transient expression of proteins, we used pCAGGS::EFF-1-V5 (Avinoam et al., 2011), pCAGGS::AtHAP2-V5, or pCAGGS::AtHAP2-YFP.

To label nuclei and cytoplasm in mixing assays, we used peGFP-N1 (Addgene plasmid 6085-1; Takara Bio Inc.) provided by M. Haus-Cohen and Y. Reiter (Technion-Israel Institute of Technology, Haifa, Israel). To label the cytoplasm, we used pRFPnes (pRFPcyto; Hu et al., 2003), provided by C. Giraudo and J. Rothman (Yale University, New Haven, CT). Oligonucleotides were synthesized by Syntezza Bioscience, restriction enzymes were purchased from New England Bi-olabs, Inc., or Thermo Fisher Scientific, and ligations were performed using T4 ligase (Promega).

Fusion assays

For content mixing (fusion) assay, BHK cells at 70% confluence were cotransfected (using Fugene HD; Promega) with 2 µg pGENE B::EFF-1-V5 (Pérez-Vargas et al., 2014) or pGENE B::AtHAP2-V5 and 1 µg pSWITCH to obtain BHK-EFF-1 or BHK-HAP2 cells. As cotransfection markers, we used RFP containing a nuclear export signal (pRFPcyto; Avinoam et al., 2011) or peGFP-N1 (Takara Bio Inc.). Control cells were transfected with pRFPcyto or peGFP-N1. 4 h after transfection, the cells were washed twice with PBS⁻ in the plates (35-mm Nunclon Surface) and detached using 0.05% EDTA solution (Biological Industries). BHK cells were collected in Eppendorf tubes, resuspended in DMEM with 10% FBS, and counted; equal amounts of cells were mixed and seeded in glass-bottom plates (12-well black, glass-bottom #1.5H; Cellvis). 4 h after mixing the EFF-1 or AtHAP2, expression vectors were induced, changing the medium with DMEM

containing 10% FBS and 10^{-7} M mifepristone. After 18 h, the cells were fixed with 4% PFA in PBS and processed for immunofluorescence (see Immunofluorescence section). To assay multinucleated cells and detect the transfected proteins (AtHAP2-V5 or EFF-1-V5), we stained cells with anti-V5 mAb (Invitrogen) and nuclei with 1 µg/ml DAPI (Podbilewicz et al., 2006). Micrographs were obtained using wide-field laser illumination using an ELYRA system S.1 microscope at a magnification of 20× (Plan-Apochromat 20×/NA 0.8; ZEISS) or 25 (LCI Plan-Neofluar 25×/NA 0.8; ZEISS). We counted the number of nuclei in cells expressing RFPcyto or GFP. For cytoplasmic mixing, we counted the number of GFP nuclei in RFPcyto-positive cells. Transfection efficiency was evaluated as 13–16% based on RFPcyto, GFP, and anti-V5 immunofluorescence. The fusion indexes (shown as percentage of multinucleation) were defined as the ratio between the number of nuclei in multinucleated cells (Nm) and the total number of nuclei in multinucleated cells and expressing cells that were in contact (Nc) but did not fuse using the following equation: % multinucleation = (Nm/Nc + Nm) × 100. The percentage of multinucleation was calculated independently for GFP- or RFPcyto-expressing cells. DAPI was used to confirm the number of nuclei in expressing cells. GFP (also green nuclei) and RFPcyto (nuclei not red; "black").

The GFP + RFP mixing index was calculated using the following equation: % mixing = (Ngfp + rfp/nCrfp + gfp) × 100. Ngfp + rfp is the number of GFP nuclei in cells with cytoplasmic RFP. nCrfp + gfp is the number of contacts between RFPcyto- and GFP-expressing cells. Note that RFPcyto or GFP mononucleated cells that were not in contact in the micrographs were not counted. The multinucleation and mixing indexes are presented as means ± SEM of three independent experiments ($n \geq 400$ for RFPcyto, GFP, and RFPcyto + GFP calculations in each experiment). The images were obtained using wide-field microscopy with laser illumination in an ELYRA system S.1 microscope at a magnification of 63 (Plan-Apochromat NA 1.4).

Hemifusion assay

BHK cells were grown in 35-mm tissue-culture dishes to 75% confluence and transfected to express HAP2 as described in the Cells and reagents section. 15 h later, the cells were labeled with fluorescent lipid Vybrant DiI and CellTracker green (V22885 and C7025; Thermo Fisher Scientific) at final concentrations of 4 and 1 µM, respectively (donor cells). 1 h later (at 16 h after transfection), the cells were co-plated with either unlabeled transfected cells or unlabeled nontransfected cells at a 1:1 ratio (acceptor cells). The cells were fixed with phosphate-buffered 4% formaldehyde for 10 min at 22°C and, 8 h later, labeled with Hoechst 33342 and analyzed using fluorescence microscopy. To quantify hemifusion, the numbers of cells labeled only with the membrane probe in each field of view were normalized to the total number of cells labeled with both membrane and content probes. For transfected donor cells mixed with transfected acceptor cells, the total number of double-labeled cells was 1,451. For transfected donor cells that were mixed with nontransfected acceptor cells, the total number of double-labeled cells analyzed was 1,034. For nontransfected donor cells that were mixed with nontransfected acceptor cells, the total number of double-labeled cells analyzed was 1,742.

Immunofluorescence

BHK cells were grown on 35-mm tissue-culture plates with glass coverslips or in glass-bottom plates (12-well black, glass-bottom #1.5H). Nonpermeabilized cells were placed on ice and fixed with 4% PFA in PBS, washed twice with ice-cold PBS followed by incubation in 40 mM NH₄Cl to block free aldehydes, and blocked in ice-cold 1% FBS in PBS. Permeabilized cells were fixed with 4% PFA in PBS, followed by incubation in 40 mM NH₄Cl to block free aldehydes, washed

in PBS, permeabilized in 0.1% Triton X-100 in PBS, and blocked in 1% FBS in PBS. After fixation, the plates/coverslips were incubated 1 h with mouse anti–V5 antibody (Invitrogen) 1:500 or with rabbit anti At-HAP2 1:250 in PBS. The secondary antibodies were goat anti–mouse coupled to Alexa Fluor 643 or goat anti–rabbit coupled to Alexa Fluor 568 (Molecular Probes/Invitrogen) diluted 1:500 in PBS. Nuclei were visualized with 1 µg/ml DAPI (Avinoam et al., 2011). Permeabilized and nonpermeabilized BHK cells were used as negative controls and gave background staining. The images were obtained using wide-field microscopy with laser illumination in an ELYRA system S.1 microscope at a magnification of 63 (Plan-Apochromat NA 1.4).

Live imaging of fusing cells
Time-lapse microscopy to identify fusing cells was performed using a spinning disc confocal microscope (CSU-X; Yokogawa Electric Corporation) with an Eclipse Ti and a Plan Apo 20× (NA 0.75; Nikon) objective. Six-well plates with glass bottoms (Cellvis) were incubated in an Oko laboratory CO_2- and temperature-controlled chamber at 37°C and 5% CO_2. Cells were cotransfected with pGENE B::AtHAP2-V5, pSWITCH, and pRFPcyto to obtain BHK-HAP2–expressing RFPcyto. 6–18 h postinduction, images in differential interference contrast and red channels were obtained every 2 to 3 min in different positions of the plate using high gain and minimum laser exposure. Images were captured with an iXon 3 EMCCD camera (Andor Technology). Fusing cells were identified based on mixing of cytoplasms containing RFP-cyto. Identified syncytia were imaged at higher resolution using an Apo 60× (NA 1.4) objective. Confocal z-series were obtained to confirm the formation of multinucleated giant cells. Image analyses were done in Metamorph and ImageJ (National Institutes of Health).

Cell cycle inhibition
To test whether multinucleation was because of failure of cytokinesis, we used FdUrd treatment to arrest BHK cells at G_1/S phase (Dijkwel et al., 1986). BHK-21 cells were grown on tissue-culture plates with glass coverslips (1.5H; Marienfeld). The cells at 70% confluence were cotransfected (using Fugene HD) with 2 µg/µl pCAGGS::AtHAP2 DNA and 1 µg/µl pEGFP-N1. For negative control, BHK-21 were transfected with 1 µg/µl pEGFP-N1. 4 h after transfection, the medium was changed with fresh DMEM with 10% FBS and 2 mM FdUrd (Sigma-Aldrich). 16–18 h after FdUdr addition, the cells were fixed with 4% PFA, permeabilized, stained with DAPI, and incubated with mouse anti–V5, followed by anti–mouse Alexa Fluor 568. We calculated multinucleation (see Fussion assays section) and compared the multinucleation index with control transfected cells without FdUrd. The mean results of two independent experiments are shown. $n > 800$ nuclei for each experiment.

Surface biotinylation of proteins expressed on BHK cells
BHK cells at 70% confluence were transfected with 2 µg pCAGGS, pCAGGS::EFF-1-V5, or pCAGGS::AtHAP2-V5. 18–24 h posttransfection, cells were washed twice with ice-cold PBS^{2+} and labeled with EZ-Link Sulfo NHS-Biotin (Thermo Fisher Scientific) for 30 min on ice. Surface biotinylation was followed by four washes with ice-cold PBS^{2+} and one wash with DMEM with 10% FBS to quench residual biotin followed by two washes with PBS^{2+}. After the washes, 300 µl lysis buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM EDTA, and 1% Triton X-100 supplemented with chymostatin, leupeptin, antipain, and pepstatin and 10 mM iodoacetamide) was added to each plate, and the plates were incubated 15 min on ice. After removal of insoluble debris by centrifugation (10 min at 21,000 g), the lysate was mixed with NeutrAvidin Agarose Resin (Thermo Fisher Scientific), and SDS was supplemented to a final concentration of 0.3%. Affinity puri-

fication of biotinylated proteins was performed incubating for 2–12 h at 4°C. The precipitated complex was then mixed with SDS-PAGE loading buffer with freshly added 50 mM DTT and incubated 5 min at 100°C. After pelleting by centrifugation, the samples were separated on 4–12% SDS-PAGE gel (Invitrogen) and analyzed by Western blotting using anti–V5 mouse monoclonal antibodies (1:5,000; Invitrogen) followed by 1:15,000 HRP goat anti–mouse antibodies (The Jackson Laboratory). Loading was controlled using anti-actin C4 monoclonal 1:2,000 (MP Biomedicals). Data shown are representative of at least three independent experiments (Podbilewicz et al., 2006).

Pseudoviruses
We prepared pseudoviruses based on the method originally used to analyze Ebola virus glycoprotein (Takada et al., 1997) and modified for AFF-1 and EFF-1 (Avinoam et al., 2011). BHK cells at 70% confluence were transfected with 2 µg pCAGGS::EFF-1-V5 or pCAGGS ::AtHAP2-V5. After 24 h at 37°C in 5% CO_2, cells were infected with VSVΔG-G at an MOI of 2–5 for 1 h at 37°C in 5% CO_2 in serum-free medium (DMEM). After six washes to remove unabsorbed VSVΔG-G, we incubated the cells for 24 h at 37°C. The supernatants containing VSVΔG-HAP2 and VSVΔG-EFF-1 were collected and centrifuged at 600 g for 10 min at 4°C to clear cell debris. Virions were collected and concentrated by two consecutive centrifugations at 100,000 g through a 20 and 10% sucrose cushion and resuspended in 130 mM NaCl and 25 mM Hepes, pH 7.4.

Determination of the titers of VSV pseudotype viruses
To determine the titer of each pseudovirus preparation, 10^4 BHK cells were plated in each well of a 96-well tissue-culture plate. BHK cells were transfected with 1 µg/ml DNA pCAGGS::EFF-1-V5 or pCAG GS::HAP2-V5 or nontransfected. We performed six serial dilutions of the viruses that were used to inoculate the cells. After 18–24 h of incubation, GFP-expressing cells were counted in at least two dilutions using an ELYRA system S.1 microscope at a magnification of 20 (Plan-Apochromat 20×/NA 0.8). Each experiment was repeated at least three times. All of the infections were done in the presence of anti-G mAb to prevent the activity of any residual VSVΔG-G (Avinoam et al., 2011).

Data analysis
We did not predetermine sample size using statistical methods. The experiments were not randomized. We were not blinded to allocation during experiments and assessment of outcomes. Interobserver error was estimated for counting of multinucleated cells, cells in contact, and content-mixing experiments. To estimate interobserver variation, two investigators counted cells in contact, multinucleation, and mixing of contents for the same 20 micrographs of different fields (examples of the micrographs used can be seen in Fig. 3 B). The differences between percentages of multinucleation and content mixing obtained by the two observers was <10%. Figures were prepared with Photoshop CS5 (Adobe Systems), Illustrator CS6 (Adobe Systems), and ImageJ.

Statistical tests
The results are expressed as means ± SEM. For each experiment, at least three independent biological repetitions were performed. We evaluated the significance of differences between mean values by using the unpaired t test function (GraphPad Software).

Online supplemental material
Fig. S1 shows confocal sections of fused HAP2-BHK cells and low surface localization of AtHAP2 in mammalian cells. Fig. S2 shows graphic summaries of HHblits and LOMETS structural prediction results for proteins of the HAP2/GCS1 family. Fig. S3 shows global structural

comparisons and HMMs profile-to-profile analyses of Fusexins. Videos 1–4 show time-lapse imaging of BHK-HAP2 cells undergoing cell fusion. Table S1 shows that FdUrd treatment increases HAP2-mediated cell fusion. Table S2 is a dataset including HHblits and LOMETS structural prediction results. Data file 1 (ITASSERectodomains.pdb) contains the structures shown in Fig. S3 C. Data file 2 (profilePhylogeny.py) contains the automation script for generation of profile-based phylogenies using HHblits and HHsearch. Data file 3 (StructuresToMSAandPhylo.py) contains the automation script for the generation of structure-based phylogenies using TMalign and FATCAT.

References

Aguilar, P.S., M.K. Baylies, A. Fleissner, L. Helming, N. Inoue, B. Podbilewicz, H. Wang, and M. Wong. 2013. Genetic basis of cell-cell fusion mechanisms. *Trends Genet.* 29:427–437. http://dx.doi.org/10.1016/j.tig.2013.01.011

Avinoam, O., K. Fridman, C. Valansi, I. Abutbul, T. Zeev-Ben-Mordehai, U.E. Maurer, A. Sapir, D. Danino, K. Grünewald, J.M. White, and B. Podbilewicz. 2011. Conserved eukaryotic fusogens can fuse viral envelopes to cells. *Science.* 332:589–592. http://dx.doi.org/10.1126/science.1202333

Baker, N.A., D. Sept, S. Joseph, M.J. Holst, and J.A. McCammon. 2001. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA.* 98:10037–10041. http://dx.doi.org/10.1073/pnas.181342398

Bianchi, E., B. Doe, D. Goulding, and G.J. Wright. 2014. Juno is the egg Izumo receptor and is essential for mammalian fertilization. *Nature.* 508:483–487. http://dx.doi.org/10.1038/nature13203

Chernomordik, L.V., and M.M. Kozlov. 2005. Membrane hemifusion: crossing a chasm in two leaps. *Cell.* 123:375–382. http://dx.doi.org/10.1016/j.cell.2005.10.015

Cole, E.S., D. Cassidy-Hanley, J. Fricke Pinello, H. Zeng, M. Hsueh, D. Kolbin, C. Ozzello, T. Giddings Jr., M. Winey, and T.G. Clark. 2014. Function of the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate. *Curr. Biol.* 24:2168–2173. http://dx.doi.org/10.1016/j.cub.2014.07.064

Dijkwel, P.A., P.W. Wenink, and J. Poddighe. 1986. Permanent attachment of replication origins to the nuclear matrix in BHK-cells. *Nucleic Acids Res.* 14:3241–3249. http://dx.doi.org/10.1093/nar/14.8.3241

Dolinsky, T.J., P. Czodrowski, H. Li, J.E. Nielsen, J.H. Jensen, G. Klebe, and N.A. Baker. 2007. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* 35:W522–W525. http://dx.doi.org/10.1093/nar/gkm276

Forslund, K., A. Henricson, V. Hollich, and E.L. Sonnhammer. 2008. Domain tree-based analysis of protein architecture evolution. *Mol. Biol. Evol.* 25:254–264. http://dx.doi.org/10.1093/molbev/msm254

Gattegno, T., A. Mittal, C. Valansi, K.C. Nguyen, D.H. Hall, L.V. Chernomordik, and B. Podbilewicz. 2007. Genetic control of fusion pore expansion in the epidermis of *Caenorhabditis elegans*. *Mol. Biol. Cell.* 18:1153–1166. http://dx.doi.org/10.1091/mbc.E06-09-0855

Harrison, S.C. 2008. Viral membrane fusion. *Nat. Struct. Mol. Biol.* 15:690–698. http://dx.doi.org/10.1038/nsmb.1456

Holm, L., and P. Rosenström. 2010. Dali server: Conservation mapping in 3D. *Nucleic Acids Res.* 38:W545–W549. http://dx.doi.org/10.1093/nar/gkq366

Holm, L., S. Kääriäinen, P. Rosenström, and A. Schenkel. 2008. Searching protein structure databases with DaliLite v.3. *Bioinformatics.* 24:2780–2781. http://dx.doi.org/10.1093/bioinformatics/btn507

Hu, C., M. Ahmed, T.J. Melia, T.H. Söllner, T. Mayer, and J.E. Rothman. 2003. Fusion of cells by flipped SNAREs. *Science.* 300:1745–1749. http://dx.doi.org/10.1126/science.1084909

Igonet, S., and F.A. Rey. 2012. SnapShot: Viral and eukaryotic protein fusogens. *Cell.* 151:1634–1634.e1. http://dx.doi.org/10.1016/j.cell.2012.11.041

Johnson, M.A., K. von Besser, Q. Zhou, E. Smith, G. Aux, D. Patton, J.Z. Levin, and D. Preuss. 2004. *Arabidopsis* hapless mutations define essential gametophytic functions. *Genetics.* 168:971–982. http://dx.doi.org/10.1534/genetics.104.029447

Kalev, I., M. Mechelke, K.O. Kopec, T. Holder, S. Carstens, and M. Habeck. 2012. CSB: A Python framework for structural bioinformatics. *Bioinformatics.* 28:2996–2997. http://dx.doi.org/10.1093/bioinformatics/bts538

Kielian, M., C. Chanel-Vos, and M. Liao. 2010. Alphavirus entry and membrane fusion. *Viruses.* 2:796–825. http://dx.doi.org/10.3390/v2040796

Lefort, V., R. Desper, and O. Gascuel. 2015. FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol. Biol. Evol.* 32:2798–2800. http://dx.doi.org/10.1093/molbev/msv150

Liu, Y., R. Tewari, J. Ning, A.M. Blagborough, S. Garbom, J. Pei, N.V. Grishin, R.E. Steele, R.E. Sinden, W.J. Snell, and O. Billker. 2008. The conserved plant sterility gene HAP2 functions after attachment of fusogenic membranes in *Chlamydomonas* and *Plasmodium* gametes. *Genes Dev.* 22:1051–1068. http://dx.doi.org/10.1101/gad.1656508

Liu, Y., J. Pei, N. Grishin, and W.J. Snell. 2015. The cytoplasmic domain of the gamete membrane fusion protein HAP2 targets the protein to the fusion site in *Chlamydomonas* and regulates the fusion reaction. *Development.* 142:962–971. http://dx.doi.org/10.1242/dev.118844

Maruyama, D., M. Ohtsu, and T. Higashiyama. 2016. Cell fusion and nuclear fusion in plants. *Semin. Cell Dev. Biol.* 60:127–135. http://dx.doi.org/10.1016/j.semcdb.2016.07.024

Mori, T., H. Kuroiwa, T. Higashiyama, and T. Kuroiwa. 2006. GENERATIVE CELL SPECIFIC 1 is essential for angiosperm fertilization. *Nat. Cell Biol.* 8:64–71. http://dx.doi.org/10.1038/ncb1345

Okamoto, M., L. Yamada, Y. Fujisaki, G. Bloomfield, K. Yoshida, H. Kuwayama, H. Sawada, T. Mori, and H. Urushihara. 2016. Two HAP2-GCS1 homologs responsible for gamete interactions in the cellular slime mold with multiple mating types: Implication for common mechanisms of sexual reproduction shared by plants and protozoa and for male-female differentiation. *Dev. Biol.* 415:6–13. http://dx.doi.org/10.1016/j.ydbio.2016.05.018

Pérez-Vargas, J., T. Krey, C. Valansi, O. Avinoam, A. Haouz, M. Jamin, H. Raveh-Barak, B. Podbilewicz, and F.A. Rey. 2014. Structural basis of eukaryotic cell-cell fusion. *Cell.* 157:407–419. http://dx.doi.org/10.1016/j.cell.2014.02.020

Podbilewicz, B. 2014. Virus and cell fusion mechanisms. *Annu. Rev. Cell Dev. Biol.* 30:111–139. http://dx.doi.org/10.1146/annurev-cellbio-101512-122422

Podbilewicz, B., E. Leikina, A. Sapir, C. Valansi, M. Suissa, G. Shemer, and L.V. Chernomordik. 2006. The *C. elegans* developmental fusogen EFF-1 mediates homotypic fusion in heterologous cells and in vivo. *Dev. Cell.* 11:471–481. http://dx.doi.org/10.1016/j.devcel.2006.09.004

Remmert, M., A. Biegert, A. Hauser, and J. Söding. 2012. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods.* 9:173–175. http://dx.doi.org/10.1038/nmeth.1818

Sapir, A., J. Choi, E. Leikina, O. Avinoam, C. Valansi, L.V. Chernomordik, A.P. Newman, and B. Podbilewicz. 2007. AFF-1, a FOS-1-regulated fusogen, mediates fusion of the anchor cell in *C. elegans*. *Dev. Cell.* 12:683–698. http://dx.doi.org/10.1016/j.devcel.2007.03.003

Sievers, F., A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. http://dx.doi.org/10.1038/msb.2011.75

Söding, J. 2005. Protein homology detection by HMM-HMM comparison. *Bioinformatics.* 21:951–960. http://dx.doi.org/10.1093/bioinformatics/bti125

Takada, A., C. Robison, H. Goto, A. Sanchez, K.G. Murti, M.A. Whitt, and Y. Kawaoka. 1997. A system for functional analysis of Ebola virus glycoprotein. *Proc. Natl. Acad. Sci. USA.* 94:14764–14769. http://dx.doi.org/10.1073/pnas.94.26.14764

Tsirigos, K.D., C. Peters, N. Shu, L. Käll, and A. Elofsson. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res.* 43:W401–W407. http://dx.doi.org/10.1093/nar/gkv485

von Besser, K., A.C. Frank, M.A. Johnson, and D. Preuss. 2006. Arabidopsis HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. *Development.* 133:4761–4769. http://dx.doi.org/10.1242/dev.02683

Wang, Y., B.W. O'Malley Jr., S.Y. Tsai, and B.W. O'Malley. 1994. A regulatory system for use in gene transfer. *Proc. Natl. Acad. Sci. USA.* 91:8180–8184. http://dx.doi.org/10.1073/pnas.91.17.8180

Waterhouse, A.M., J.B. Procter, D.M. Martin, M. Clamp, and G.J. Barton. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* 25:1189–1191. http://dx.doi.org/10.1093/bioinformatics/btp033

White, J.M., S.E. Delos, M. Brecher, and K. Schornberg. 2008. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit. Rev. Biochem. Mol. Biol.* 43:189–219. http://dx.doi.org/10.1080/10409230802058320

Wong, J.L., and M.A. Johnson. 2010. Is HAP2-GCS1 an ancestral gamete fusogen? *Trends Cell Biol.* 20:134–141. http://dx.doi.org/10.1016/j.tcb.2009.12.007

Wong, J.L., A.R. Leydon, and M.A. Johnson. 2010. HAP2(GCS1)-dependent gamete fusion requires a positively charged carboxy-terminal domain. *PLoS Genet.* 6:e1000882. http://dx.doi.org/10.1371/journal.pgen.1000882

Wu, S., and Y. Zhang. 2007. LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic Acids Res.* 35:3375–3382. http://dx.doi.org/10.1093/nar/gkm251

Wu, S., and Y. Zhang. 2008. MUSTER: Improving protein sequence profile-profile alignments by using multiple sources of structure information. *Proteins.* 72:547–556. http://dx.doi.org/10.1002/prot.21945

Wu, S., J. Skolnick, and Y. Zhang. 2007. Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biol.* 5:17. http://dx.doi.org/10.1186/1741-7007-5-17

Xu, J., and Y. Zhang. 2010. How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics.* 26:889–895. http://dx.doi.org/10.1093/bioinformatics/btq066

Yang, J., R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang. 2015. The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods.* 12:7–8. http://dx.doi.org/10.1038/nmeth.3213

Ye, Y., and A. Godzik. 2004. FATCAT: A web server for flexible structure comparison and structure similarity searching. *Nucleic Acids Res.* 32:W582–W585. http://dx.doi.org/10.1093/nar/gkh430

Zhang, Y., and J. Skolnick. 2005. TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* 33:2302–2309. http://dx.doi.org/10.1093/nar/gki524