Journal of Genetics and Genomics 45 (2018) 681–684

Contents lists available at ScienceDirect

Journal of Genetics and Genomics

Journal homepage: www.journals.elsevier.com/journal-of-genetics-and-genomics/

Letter to the editor

ZFN, TALEN and CRISPR-Cas9 mediated homology directed gene insertion in *Arabidopsis*: A disconnect between somatic and germinal cells

Breakthroughs in the generation of programmable sequence-specific nucleases (SSNs), such as zinc finger nucleases (ZFNs), TAL effector nucleases (TALENs) and the RNA-directed nuclease CRISPR-associated protein 9 (Cas9), have greatly increased the ease of plant genome engineering (Voytas, 2013; Malzahn et al., 2017). Programmable SSNs introduce a DNA double-strand break (DSB) at a target site in the genome that must be repaired by one of several endogenous DNA repair pathways. Non-homologous end joining (NHEJ) and microhomology-mediated end joining (MMEJ) may result in small insertions and deletions, while homologous recombination (HR), a template-driven repair, is capable of copying information from a homologous donor molecule co-delivered with the nuclease. This precise HR-based modification or gene targeting (GT) remains challenging in most plants (Malzahn et al., 2017). Factors limiting plant GT frequencies include the low frequency that HR is used for DNA repair and the difficulty in delivering abundant donor templates to plant cells. GT in *Arabidopsis thaliana* has proven particularly difficult to optimize; several recent publications have detailed efforts to enhance GT efficiencies but were unable to improve upon existing methods (Shaked et al., 2005; de Pater et al., 2013; Schiml et al., 2014). Here we present data detailing a disconnect in *Arabidopsis* GT efficiencies between somatic and germinal tissues, our attempts to close that gap, and pitfalls of relying on somatic GT reporters.

We first tested ZFN-mediated GT of a *RLK* gene (*At1g53430*). We constructed a T-DNA plasmid containing an estrogen inducible ZFN pair targeting *RLK*, followed by a donor molecule that would introduce an in-frame *BAR* gene (Fig. S1A). Notably, the ZFN also targets a homologous gene (*At1g53440*) 10 kb downstream of *At1g53430*. We predicted that in stably transformed *Arabidopsis* plants, inducible expression of the ZFN should facilitate HR, resulting in expression of *BAR* from the promoter of *At1g53430* (Fig. S1B). T_1 transgenic plants were screened on MS medium supplied with hygromycin (for the transgene) and β-estradiol (to induce ZFN expression) (Fig. S1C). We detected ZFN-induced chromosomal deletions by PCR, confirming that ZFN expression was effectively induced by β-estradiol (Fig. S1D). After transferring transgenic plants from plates to soil, ZFN expression was continuously induced by spraying β-estradiol (10 μM) every two days for approximately one month. Then, T_2 seeds from 50 independent T_1 parental lines were collected and screened on MS medium containing BASTA, and two T_1 lines (#22 and #32) yielded BASTA-resistant plants. HR events were detected by PCR in these plants, consistent with the BASTA-resistant phenotype (Fig. S1E). After following four promising T_2 lines (#22-1-#22-4) to T_3 generation, however, only very few BASTA-resistant plants were obtained in the T_3 generation, suggesting that the HR events in these T_2 lines are somatic events. Our unsuccessful attempt to achieve HR with *BAR* as a selectable marker is consistent with a recent report (Hahn et al., 2017).

Our next strategy to detect HR used a β-glucuronidase (GUS) reporter and constitutive promoters for SSN expression. Given that ZFNs may display toxicity when constitutively expressed, we used TALENs to create targeted DNA DSBs. Three constitutive promoters (*AtUBQ10*, *AtUBQ1*, and 2× *35S*) were evaluated for their ability to drive YFP reporter or TALEN expression in tobacco protoplasts (Figs. S2A, S2B, S3 and S4), and TALEN expression in stable transgenic *Arabidopsis* lines (Figs. S2C and S5). *AtUBQ10* performed best and was selected for all TALEN GT experiments. A T-DNA plasmid (based on pYPQ202) containing a TALEN pair targeting *NATA2* and a donor molecule encoding a 3′ GUS fusion was delivered to *ku70 Arabidopsis* plants, which are deficient in NHEJ repair and thereby have increased HR activity (Christian et al., 2013; Qi et al., 2013) (Figs. 1A and S6). The *GUS* transgene was separated from the *NATA2* gene by a self-cleaving T2A sequence. 101 T_1 plants were assessed for somatic GT events with GUS staining. Seven had high levels of staining and were propagated for further analysis. Three of those seven lines (#26, #57, and #87) produced GUS-positive T_2 seedlings and, notably, line #87 produced 27 of 100 progenies that stained entirely blue (Figs. 1B and S7). To determine if high levels of somatic GT serve as a proxy for germinal GT, 15 T_2 plants from line #87 were screened for GUS activity to determine if the staining was heritable. Five of the 15 plants had entirely blue leaves, a near-Mendelian segregation suggesting that germinal GT had been achieved. Five T_3 progeny of each of these lines were further screened. The results showed that nearly all sampled leaves displayed the GT phenotype and GT-specific PCR bands were confirmed by Sanger sequencing (Figs. S8 and S9). Despite this strong phenotypic evidence, Southern blot did not confirm germinal modification, and instead demonstrated an inability to separate the GT genotype from the T-DNA (Figs. 1C and S10). By tracking the most promising somatic GT events, we appeared to have selected highly active somatic GT lines, confounding attempts to identify germinal GT events.

Due to a dearth of TALEN-mediated GT in *Arabidopsis*, we sought to determine if nuclease choice could be partially responsible for incongruities in somatic and germinal events observed above. Additionally, we sought to enrich GT events in the germline cells.

https://doi.org/10.1016/j.jgg.2018.07.011

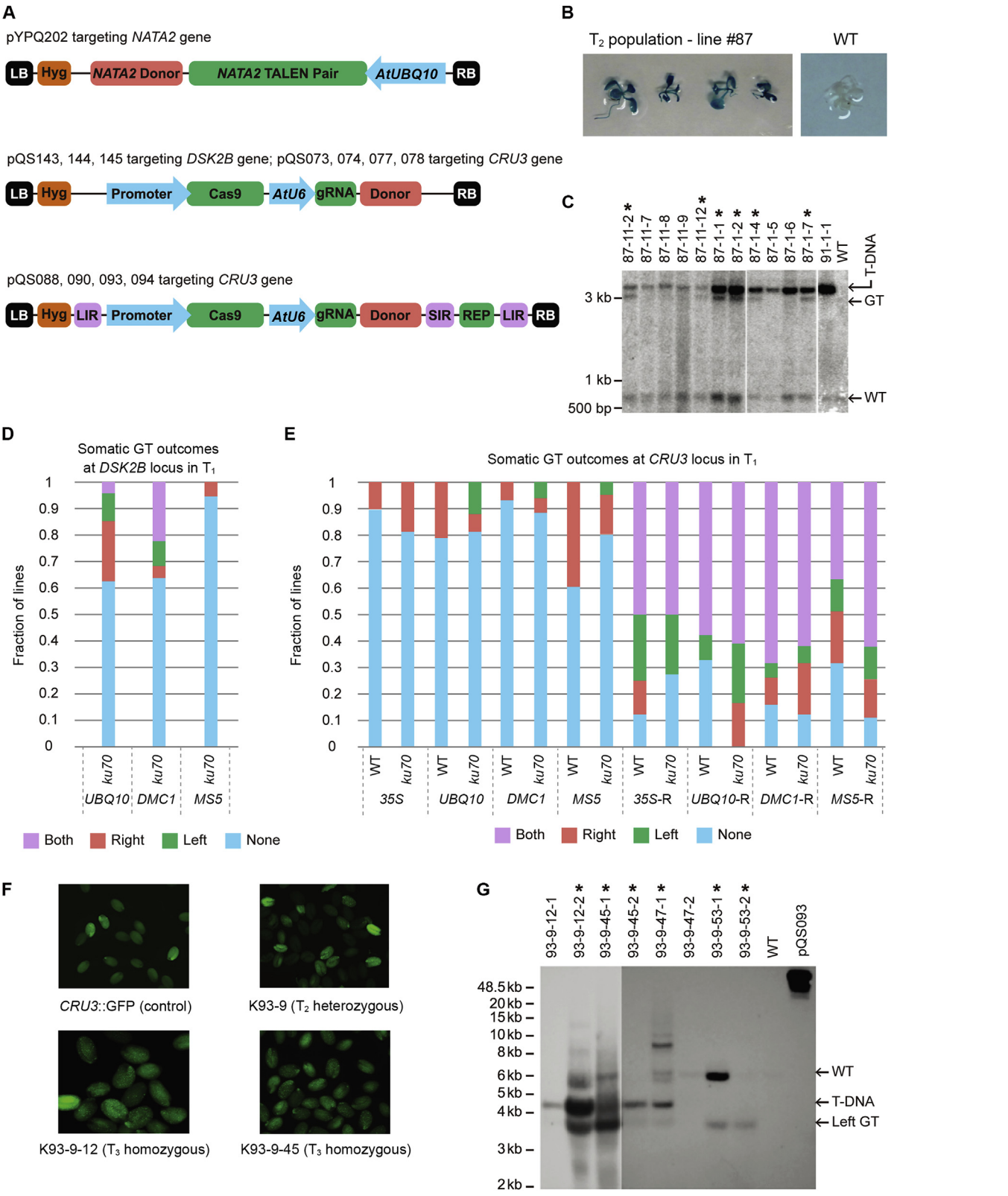

Fig. 1. TALEN and CRISPR-Cas9 mediate gene targeting (GT) at the *NATA2*, *DSK2B* and *CRU3* loci in *Arabidopsis*. **A**: Schematic shows the GT constructs targeting the *NATA2*, *DSK2B* and *CRU3* genes with TALEN or CRISPR-Cas9. **B**: GUS staining of *Arabidopsis* seedlings. Four of the 27 whole blue T_2 seedlings from the #87 T_1 line were shown here. The wild-type (WT) plant served as a control. **C**: Southern blot analysis at the *NATA2* locus with selected T_3 lines. Six T_3 lines from the #87-1 T_2 parent and five T_3 lines from the #87-11 T_2 parent were evaluated. An independent transgenic T_3 line (#91-1-1) and WT plants were used as controls. The three bands corresponding to the T-DNA, HR/GT event and unmodified locus (WT) were indicated by arrows. Plants having undergone HR were marked with asterisks. **D**: Column shows the somatic GT outcomes in T_1 generation at the *DSK2B* locus. The GT events were quantified by a PCR based method. Transgenic plants were in *ku70* mutant background. *UBQ10*, *DMC1* and *MS5* stand for *Cas9* expression under the promoter in non-GVR vectors, respectively. Sample size for each column is 21, 22 and 16 (left to right). **E**: Column shows the somatic GT outcomes in T_1 generation at the *CRU3* locus. The GT events were quantified by PCR based method. Transgenic plants were in Columbia (WT) or *ku70* mutant background. *35S*-R, *UBQ10*-R, *DMC1*-R and *MS5*-R stand for *Cas9* expression under

To this end, we selected CRISPR-Cas9 and meiosis-specific promoters for further GT experiments. The gRNA targeting the second exon of *DSK2B* was generated and a donor molecule was designed to insert *T2A-GUS* at the 3′ end of the gene (Figs. 1A, S11A and S12). The *DSK2B* gRNA and donor were delivered to *ku70 Arabidopsis* plants with *Cas9* under the control of *AtUBQ10* (a constitutive promoter), or *DMC1* (a meiosis-specific promoter), or *MS5* (a meiosis-specific promoter). 59 T_1 lines containing one of the three constructs were screened by PCR for the presence of both GT junctions (Figs. 1D and S13). Leaves from 59 T_1 lines were stained for GUS activity, and two lines (#143-13 and #144-16) were found to be GUS-positive: line #144-16 was found by PCR to have both GT junctions, and line #143-13 had only the left GT junction (Table S1). Additionally, five other lines (#143-5, #144-2, #144-4, #144-5 and #144-8) were positive for GT via PCR but negative for GUS activity. This could be explained by higher sensitivity of PCR over GUS staining, or that different leaves were sampled for PCR and GUS staining. T_2 offspring of these seven lines were then GUS stained to see if the modifications were transmitted to the following generation, while only one line (#143-13) produced two GUS-positive seedlings of 86 progenies (Table S1). T_3 offspring of those GUS-positive plants all stained negatively, again indicating that following promising levels of somatic GT activity does not necessarily indicate that a germinal event has occurred. Another line (#145-3) that was negative for both GUS staining and GT junctions by PCR was found to produce a substantial portion of GUS-positive T_2 seedlings (22 of 96) (Fig. S14A and B; Table S1). Seven of these GUS-positive T_2 plants were propagated and their T_3 offspring were stained; all T_3 seedlings were found to be GUS-positive (data not shown). When further analyzed by Southern blot, these lines were again found to contain only somatic events despite the Cas9 being driven by a meiotic promoter (Fig. S14C and D).

We next sought to minimize reliance on low-throughput reporters by using a seed-specific GT phenotype (Shaked et al., 2005). To this end, eight constructs were generated, all of which contained *Cas9*, an *AtU6*-driven gRNA targeting *CRU3* (a seed storage protein) and a donor molecule encoding a 3′ GFP translational fusion that will result in GFP-positive seeds after GT (Figs. S11B and S15). These constructs vary in their *Cas9* promoter and the presence/absence of a geminiviral replicon (GVR) (Baltes et al., 2014; Cermak et al., 2015; Gil-Humanes et al., 2017). Two promoters are constitutive (*AtUBQ10* and 2× *35S*) and two are meiotic (*DMC1* and *MS5*) (Figs. 1A and S15). All eight constructs were delivered to both WT and *ku70 Arabidopsis* plants via floral dip, and 285 T_1 plants were assayed for GT events using two GT-specific PCR pairs, each detecting one side of the GT junctions (Figs. 1E and S16). We clearly saw an impact of the replicating GVR on frequencies of somatic GT: no sampled non-GVR leaves contained both junctions while both were detectable in 50% of GVR samples (Figs. S17, S18 and S19). Unexpectedly, the promoter choice also appeared to have little impact on levels of somatic mutagenesis. Both meiotic promoters showed similar frequencies/patterns of GT when compared with the constitutive promoters, with and without GVR (Fig. 1E).

We then screened the seeds of the T_1 lines for GFP fluorescence. None of the 125 non-GVR T-DNA lines generated GFP-positive seeds, and only one of 160 GVR lines generated GFP-positive seeds (Fig. 1F). This line (#K93-9), expressing *DMC1*::Cas9, was further analyzed to determine the nature of the GT events. Initial PCR characterization indicated that it was a one-sided GT event, possibly due to repair by synthesis-dependent strand annealing (SDSA) (Fig. S16B). We hypothesized that during this SDSA event, information began to be copied near the site of the break and proceeded past the end of the distal donor arm into the T-DNA vector, which was then incorporated into the target site (Puchta and Fauser, 2013). To determine if this was the case, PCR specific to insertion events was performed and it was found that a vector sequence of at least 1.3 kb had been inserted into the target side (Fig. S20). Next, we confirmed this one-sided event is inserted at the expected genomic locus and that it is not a modified T-DNA insertion or ectopic event. The modified *CRU3* locus can be segregated away from the T-DNA, suggesting a single copy (Fig. S21A), and Southern blot demonstrated the expected band shift from the inserted sequence at the target site (Figs. 1G, S21B, S21C, and S22). Again, an extreme disconnect was found between somatic and germinal events, and that the ability of the GVR to massively increase somatic GT (from 0% detection of both GT junctions to ~57% of samples containing both GT junctions) did not translate to an increase in the number of germinal GT events.

In summary, we tested in *Arabidopsis* the ability of ZFNs, TALENs and CRISPR-Cas9, expressed from inducible, constitutive and meiotic promoters, to knock-in reporter genes (*BAR*, *GUS* and *GFP*) at endogenous loci. This work clearly demonstrates that somatic phenotypes for optimizing GT in *Arabidopsis* are unreliable; simply following lines with high levels of somatic GT will not highly enrich for germinal events. Only when utilizing a reporter that required highly tissue-specific expression, namely *GFP* targeted to *CRU3*, were we not misled by somatic events and able to capture a germinal, albeit one-sided, GT event. If the drastic disconnect in HR frequencies between somatic cells and germline cells also holds true in other species, the best practice for achieving HR in plants at present might be tissue culture, rather than directly editing germline cells *in planta*.

Acknowledgments

This work was supported by a Collaborative Funding Grant from North Carolina Biotechnology Center and Syngenta Biotechnology (2016-CFG-8003), and startup funds provided by East Carolina University and University of Maryland to Y.Q. and a grant from the National Science Foundation (IOS-1339209) to D.F.V.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jgg.2018.07.011.

References

Baltes, N.J., Gil-Humanes, J., Cermak, T., Atkins, P.A., Voytas, D.F., 2014. DNA replicons for plant genome engineering. Plant Cell 26, 151–163.
Cermak, T., Baltes, N.J., Cegan, R., Zhang, Y., Voytas, D.F., 2015. High-frequency, precise modification of the tomato genome. Genome Biol. 16, 232.
Christian, M., Qi, Y., Zhang, Y., Voytas, D.F., 2013. Targeted mutagenesis of *Arabidopsis thaliana* using engineered TAL effector nucleases (TALENs). G3 3, 1697–1705.
de Pater, S., Pinas, J.E., Hooykaas, P.J.J., van der Zaal, B.J., 2013. ZFN-mediated gene targeting of the *Arabidopsis protoporphyrinogen oxidase* gene through *Agrobacterium*-mediated floral dip transformation. Plant Biotechnol. J. 11, 510–515.
Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C.V., Sanchez-Leon, S., Baltes, N.J., Starker, C., Barro, F., Gao, C., Voytas, D.F., 2017. High-efficiency

the promoter in GVR vectors, respectively. Sample size for each column is 19, 11, 14, 16, 14, 16, 15, 20, 16, 14, 21, 18, 19, 16, 29 and 27 (left to right). **F**: Seed phenotype of *CRU3*::GFP positive control and K93-9 (T_2), K93-9-12 (T_3) and K93-9-45 (T_3). **G**: Southern blot analysis at the *CRU3* locus with selected T_3 lines. Two T_3 lines each from the T_2 parents K93-9-12, K93-9-45, K93-9-47 and K93-9-53 were evaluated. WT plants and plasmid pQS093 were used as controls. Arrows indicate the WT, T-DNA, and left GT, respectively. Plants having undergone left-side GT were marked with asterisks.

gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. Plant J. 89, 1251–1262.
Hahn, F., Mantegazza, O., Greiner, A., Hegemann, P., Eisenhut, M., Weber, A.P., 2017. An efficient visual screen for CRISPR/Cas9 activity in *Arabidopsis thaliana*. Front. Plant Sci. 8, 39.
Malzahn, A., Lowder, L., Qi, Y., 2017. Plant genome editing with TALEN and CRISPR. Cell Biosci. 7, 21.
Puchta, H., Fauser, F., 2013. Gene targeting in plants: 25 years later. Int. J. Dev. Biol. 57, 629–637.
Qi, Y., Zhang, Y., Zhang, F., Baller, J.A., Cleland, S.C., Ryu, Y., Starker, C.G., Voytas, D.F., 2013. Increasing frequencies of site-specific mutagenesis and gene targeting in *Arabidopsis* by manipulating DNA repair pathways. Genome Res. 23, 547–554.
Schiml, S., Fauser, F., Puchta, H., 2014. The CRISPR/Cas system can be used as nuclease for *in planta* gene targeting and as paired nickases for directed mutagenesis in *Arabidopsis* resulting in heritable progeny. Plant J. 80, 1139–1150.
Shaked, H., Melamed-Bessudo, C., Levy, A.A., 2005. High-frequency gene targeting in *Arabidopsis* plants expressing the yeast *RAD54* gene. Proc. Natl. Acad. Sci. U. S. A. 102, 12265–12269.
Voytas, D.F., 2013. Plant genome engineering with sequence-specific nucleases. Annu. Rev. Plant Biol. 64, 327–350.

Qiwei Shan[a], Nicholas J. Baltes[a], Paul Atkins[a], Elida R. Kirkland[b], Yong Zhang[c], Joshua A. Baller[d], Levi G. Lowder[b], Aimee A. Malzahn[e], John C. Haugner III[f,g], Burckhard Seelig[f,g], Daniel F. Voytas[a,*], Yiping Qi[b,e,h,*]

[a] *Department of Genetics, Cell Biology & Development and Center for Genome Engineering, University of Minnesota, Minneapolis, MN 55455, USA*

[b] *Department of Biology, East Carolina University, Greenville, NC 27858, USA*

[c] *Department of Biotechnology, School of Life Sciences and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China*

[d] *Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN 55455, USA*

[e] *Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD 20742, USA*

[f] *Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN 55455, USA*

[g] *BioTechnology Institute, University of Minnesota, St. Paul, MN 55108, USA*

[h] *Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, MD 20850, USA*

[*] Corresponding authors.
E-mail addresses: Voytas@umn.edu (D.F. Voytas), Yiping@umd.edu (Y. Qi).

27 April 2018
Available online 10 October 2018