ORIGINAL ARTICLE

Gene expression variability between randomly and targeted transgene integration events in tobacco suspension cell lines

Janina Kirchhoff[1] · Andreas Schiermeyer[1] · Katja Schneider[1] · Rainer Fischer[1,3] · W. Michael Ainley[2] · Steven R. Webb[2] · Helga Schinkel[1] · Stefan Schillberg[1]

Abstract

Genome editing tools such as zinc-finger nucleases provide novel strategies for genetic manipulation in plants. Unlike agrobacterium-mediated or direct gene transfer, which introduce genes randomly into the genome and thereby potentially resulting in high variation of gene expression, the targeted gene addition provides predictable integration of DNA sequences into a specified location of the plant genome. We investigated whether various independent cell lines that all contain a transgene placed in the same genomic locus by zinc-finger nuclease-mediated homologous recombination (HR) would yield a more reproducible and homogeneous level of expression compared to integration events generated via agrobacterium-mediated transformation at random sites. The variance of gene expression of targeted HR events and random integration events was analyzed in *Nicotiana tabacum* L cv. Bright Yellow 2 (BY-2) suspension cells by measuring protein amount produced from the transgene by flow cytometry, thus providing the first report on positional effects of marker gene expression in a quickly proliferating plant suspension cell line. Marker protein levels of targeted HR and single-copy random events covered a similar range; however, the uniformity of protein expression in a given cell line was significantly higher in targeted events than in lines with randomly inserted transgene; the same is true for the overall viability of protoplasts from HR lines. In conclusion, using targeted insertion into a qualified locus of a well-characterized line leads to more reliable results than random insertion into the genome.

Keywords Positional effect · Flow cytometry · Gene targeting · Homologous recombination · Tobacco BY-2 suspension cultures · Protoplasts

Introduction

In plants and plant suspension cultures, new genes or cassettes are routinely introduced for various applications in biotechnology and basic research. Most classical methods for the stable introduction of genes into plant cells rely on random integration of the transgene into the genome followed by laborious selection and screening for desired events or expression levels, since high variability of transgene expression from independent transformants is frequently observed (Butaye et al. 2005; Day et al. 2000).

Widely used plant transformation methods, such as Agrobacterium-mediated transformation or direct gene transfer like e.g. particle bombardment, do not allow a defined introduction of transgenes into the genome in terms of number or location and different levels of gene expression in independent transformants are common (Day et al. 2000). This variation in transgene expression is postulated to be attributed to several factors including chromosome location, transgene construct fidelity, and copy number, as one or multiple intact or rearranged gene copies can integrate at one or multiple loci (Hobbs et al. 1990; Kooter et al. 1999; Matzke and Matzke 1998). Furthermore, random transgene integrations

✉ Helga Schinkel
helga.schinkel@ime.fraunhofer.de

1 Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Forckenbeckstrasse 6, 52074 Aachen, Germany

2 Corteva Agriscience, 9330 Zionsville Rd, Indianapolis, IN 46268, USA

3 Present Address: Purdue University, West Lafayette, IN 47907, USA

can lead to confounding and often controversially discussed effects on expression including gene silencing and the activation or disruption of endogenous genes. Both factors cause unstable gene expression leading to unpredictable cell line behavior and potentially irreproducible results (Peach and Velten 1991).

The introduction of genes by genome editing tools such as zinc-finger nucleases (ZFN) or CRISPR/Cas9 is still a relatively novel technology. Genome editing facilitates the precise addition of transgenes to specified locations in the plant cell genome and thereby overcomes limitations of classical transformation approaches e.g. undefined loci. ZFN-mediated gene exchange and homologous recombination at defined sites using a well-characterized target cell line can save considerable time and resources for the reproducible generation of transgenic cell lines in plant research and biotechnology and is especially useful for the introduction of large cassettes to express genes for interrelated proteins to be inherited as one unit.

For mouse embryonic stem cells, it has been shown that ZFN-targeted transgene integration into the ROSA26 locus is predictable and expression is relatively uniform among targeted cell populations, whereas random plasmid/DNA integration led to highly variable levels of reporter gene expression (Perez-Pinera et al. 2012). In previous studies (Schiermeyer et al. 2019; Schneider et al. 2016), progress has been reported by applying the ZFN technology as a suitable instrument for engineering transgenic loci of tobacco BY-2 suspension cells. We established a gene landing pad (target DNA) in BY-2 target lines that comprised two markers genes (a kanamycin resistance coding gene and a gene coding for a red fluorescent protein) with internal introns allowing the targeted integration of donor DNA constructs by splitting each of the marker gene between the target and the donor construct (Schiermeyer et al. 2019). A commonly held belief is that donor DNA integration mediated by engineered nucleases would produce consistent expression of the introduced DNA. In this study, we, therefore, investigated variability in the production of the marker proteins tGFP (present on the target DNA) and DsRed (completed by the donor fragment) in our ZFN technology derived BY-2 suspension lines and lines constructed to mimic random integration of the complete corresponding tGFP and DsRed marker gene construct. From random integration events, we carefully selected suspension cell lines containing only a single complete copy located at different genome positions as well as cell lines that contained multiple copies of the cassette. Protein levels of the markers tGFP and DsRed were compared between the random lines and ZFN-mediated homologous recombination (HR) derived lines. In our experiment, fluorescence intensity of the reporter proteins, tGFP and DsRed, was relatively uniform among both homologous recombination-targeted (HRT) and randomly

integrated single-copy (SCR) cell lines, whereas multiple copy random integration (MCR) events led to highly variable levels of recombinant protein. Furthermore, we saw a greater uniformity in the different HRT lines in regard to variability of fluorescent protein production and viability compared to the SCR/MCR lines. Such uniformity makes results more predictable and reliable and is, therefore, a clear advantage of targeted insertion.

Experimental procedures

Plasmids

The target and donor vector designed for HR-mediated transgene integration is described by Schiermeyer et al. (2019). The pRandom vector (Fig. 1a) mimics a perfect integration of donor DNA into the target site by homologous recombination. This vector contains the neomycin phosphotransferase (nptII) gene (Bevan et al. 1983) interrupted by the thioredoxin-like protein gene intron from *Arabidopsis thaliana* and under the control of the nopaline synthase promoter (Pnos) (Depicker et al. 1982) followed by the DsRed coding sequence (GenBank: ADJ00077.1) under the control of a double-enhanced 35S promoter (Kay et al. 1987). The DsRed gene is interrupted by an intron sequence of the 4-coumarate CoA ligase-coding gene (At3g21230) from *A. thaliana*. As flanking marker genes the acetohydroxyacid synthase (AHAS) from *Gossypium hirsutum* (Rajasekaran et al. 1996) under the control of the ubiquitin 3 promoter from *A. thaliana* (Norris et al. 1993), conferring resistance to imazethapyr, as well as the TurboGFP gene (tGFP) from *Pontellina plumata* (Evrogen, Moscow, Russia) fused to a KDEL encoding motif for protein retrieval to the endoplasmic reticulum (Wandelt et al. 1992) under the control of the melon (*Cucumis melo*) actin 2 promoter (Clendennen et al. 2003) were used.

Plant material and culture conditions

Nicotiana tabacum L *cv.* Bright Yellow 2 (BY-2) cells were cultivated in Murashige–Skoog (MS, pH 5.8) basal medium as previously described (Kirchhoff et al. 2012; Schinkel et al. 2008). The cells were subcultured weekly by transferring 5% (v/v) of the culture into 50 mL fresh medium and maintained at 26 °C in the dark, with a constant orbital agitation of 180 rpm. Callus material from random as well as HR-targeted events was used to establish liquid cultures in 10 mL MS medium containing 1 mg/L kanamycin or 1.5 μM imazethapyr (Sigma-Aldrich, Deisenhofen, Germany) in 50-mL Tubespin Bioreactors (TPP, Trasadingen, Switzerland). After initial subculturing in the bioreactor tubes,
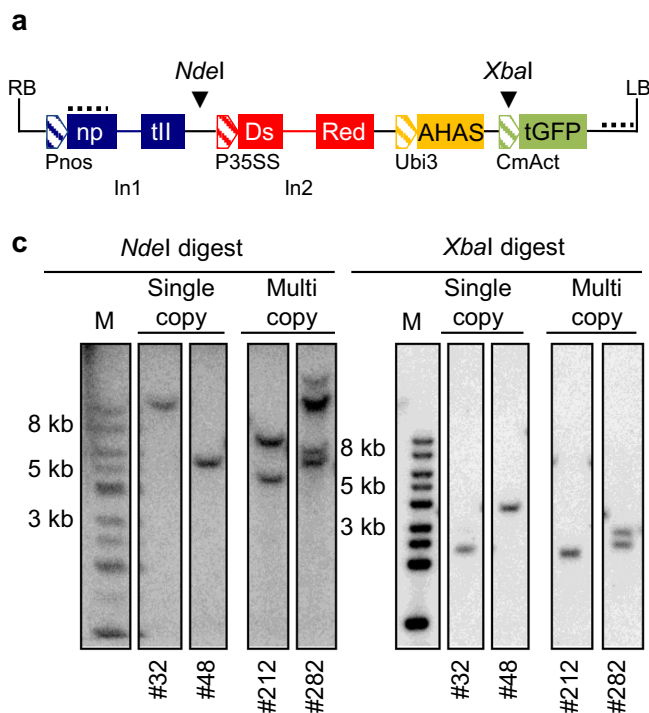
a

b

c

Analysis	Candidates / Percentage
DsRed fluorescence on imazethapyr selection	411/ 100%
Survival on kanamycin selection	407 / 99%
Single insertion by RT-qPCR	35 / 8.5%
Southern blot confirmed single target integration	20 / 4.8%
pRandom transformed true single copy events	20 / 4.8%

Fig. 1 Establishment of cell lines with randomly inserted transgene. **a** T-DNA structure of the pRandom vector with Southern blot strategy indicated. LB and RB: left and right borders; Pnos: promoter of the nopaline synthase gene; *npt*II: coding sequence of the neomycin phosphotransferase gene; In1: intron sequence of the thioredoxin-like protein gene; P35SS: promoter with duplicated enhancer of the *Cauliflower mosaic virus* 35S gene; DsRed: coding sequence for the *Discosoma* spp. red fluorescent protein containing a C-terminal KDEL motif; In2: intron sequence of the 4-coumarate CoA ligase coding gene; Ubi3: ubiquitin 3 promoter; AHAS: coding sequence of the acetohydroxyacid synthase gene; CmAct: melon actin 2 promoter; tGFP: coding sequence for the *P. plumata* green fluorescent protein

TurboGFP fused to a C-terminal KDEL encoding motif for endoplasmic reticulum retrieval. Restriction sites for genomic DNA digest indicated by black arrow heads, dashed bars represent the hybridization sites of the probes used in Southern blot analysis. **b** Summary of randomly inserted single-copy line selection. Table representing screening process of tobacco BY-2 suspension cells lines after Agrobacterium-mediated transformation. **c** Southern blot confirming randomly inserted transgene. Genomic DNA of representative samples for SCR (lines #32 and #48) and MCR (lines #212 and 282) lines, digested with *Nde*I (left panel) and *Xba*I (right panel), respectively. M: DNA size marker

cultures were subjected to routine maintenance in 50 mL in Erlenmeyer flasks with a weekly subculturing procedure.

Agrobacterium-mediated transformation and selection of single-copy random integration lines

The random vector (pRandom) was introduced into *Agrobacterium tumefaciens* strain LBA4404 (Invitrogen, Karlsruhe, Germany) by electroporation (Dower et al. 1988). Transgenic BY-2 cells were generated by co-cultivation of wild-type BY-2 cells on day four of cultivation with recombinant agrobacteria as described (An 1985). Transgenic events were selected on MS agar plates supplemented with 1.5 µM imazethapyr (Sigma-Aldrich) and screened for DsRed fluorescence by a Leica KL 1500 LCD lamp with excitation filter (BP: 545/30 nm) and foil no.182 light red (Leitz, Wetzlar, Germany). Initially resistant transformation events were tested on 1 mg/L kanamycin, and surviving callus material thereof was subjected to routine

maintenance of suspension cultures in selective MS medium. Genomic DNA was extracted from callus material using the Chemagen Plant DNA Extraction Kit (Baesweiler, Germany), and used for transgene copy number determination by qPCR and subsequent southern blot analysis. Real-time qPCR was performed using an ABI Prism 7500 Sequence Detection System (Applied Biosystems, Darmstadt, Germany) in the presence of SYBR-Green (Platinum® SYBR®Green qPCR SuperMix-UDG with ROX; Life Technologies, Darmstadt, Germany). The integration of the random construct was evaluated using Pnos and tGFP specific primer sets as well as primers amplifying the intron sequence of the internal reference gene phenylalanine ammonia-lyase (Table S1) that served as a single-copy control (Schiermeyer et al. 2019). Sample amplification was performed for 2 min at 50 °C and 10 min at 95 °C, followed by 40 cycles of 15 s 95 °C and 45 s at 60 °C. Resulting Ct values were used for a quantitative calculation of the gene copy numbers according to the 2 -ΔΔCt method (Livak and Schmittgen 2001). Extended single transgene copy

screening was done by Southern blot as described (Schneider et al. 2016), using *Nde*I (New England Biolabs, Frankfurt am Main, Germany) digested genomic DNA and a 1036 bp probe matching the *npt*II coding sequence as well as *Xba*I (New England Biolabs) digested genomic DNA in combination with a 993 bp probe binding 850 bp downstream of the tGFP gene to the T-DNA backbone.

Flow cytometric analysis

Protoplasts from BY-2 suspension cultures were isolated as described before (Schinkel et al. 2008). This was done for all three analyzed populations, i.e. HRT, SCR, and MCR lines. Qualitative analysis of the tGFP and DsRed fluorescent cell population was investigated using a FACSVerse flow cytometer (BD Bioscience, Heidelberg, Germany). The viable cell population was gated based on light scatter signals (SSC-A and FSC-H) while the percentage of fluorescent protoplasts in each culture was detected at 527/32 nm (FITC-H, GFP) and 575/26 nm (PE-H, DsRed). Three types of protoplasts were used to set the gates for the presence of red and green fluorescence and to compensate for spectral crossover: wild-type BY-2 protoplasts, protoplasts from target line TCL448 carrying the landing pad including the tGFP expression cassette (GFP+), and HRT lines where the DsRed expression cassette has been restored by the donor fragment via HR (DsRed+) (Schiermeyer et al. 2019). HRT lines used were lines examined in Schiermeyer et al. (2019) (Fig. 2, all except lines 94-1, 94-64 and 94-66). We analyzed viable gated protoplasts for each HRT ($10,400 \pm 894$), SCR ($10,169 \pm 386$) and MCR (10,000) line and processed the signal data using the FACSuite Software (BD Bioscience).

Statistics

Geometric mean values of the fluorescence measurements of each line were used to analyze the gene expression variation. Statistical analysis of the variation of the geometric mean of fluorescence for tGFP and DsRed was performed by Levenes test ($p \leq 0.05$) testing for homogeneity of variance between the three different groups of cell lines (HRT, SCR, and MCR) based on the assumption that the populations are of equal variance. Furthermore, a t-test was performed, to compare the size of standard deviations of DsRed-fluorescence in HRT lines and SCR lines to get a measure for the homogeneity of cells in one line.
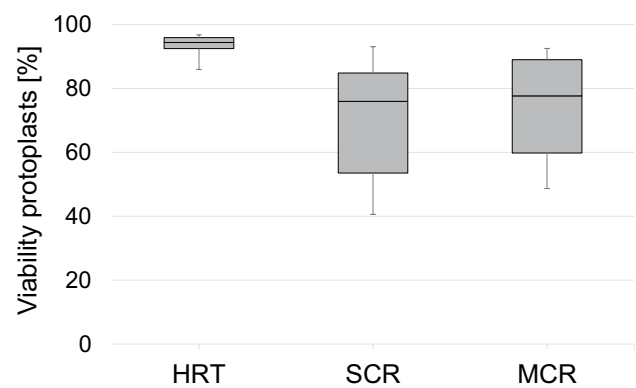

Fig. 2 Viability rate of protoplasts from different cell lines analyzed by flow cytometry. To determine the levels of DsRed and tGFP produced, viable protoplasts were gated and then used for fluorescence analysis. The percentage of protoplasts in the viability gate compared to the total number of protoplasts analyzed is displayed. 15 HR-targeted (HRT), 12 randomly inserted single-copy (SCR) cell lines and 12 randomly inserted multiple copy (MCR) cell lines were analyzed. The viability is displayed as boxplot diagrams showing the upper 75% and lower 25% percentile values (box), minimum and maximum values (whiskers) and the median/50% percentile values (line)

Results

Selection of transgenic BY-2 lines containing randomly integrated single and multiple transgene copies

To compare positional effects between targeted and random integration events, we constructed a vector (pRandom, Fig. 1a) that mimics successful donor integration into the target site in the HRT lines. In the HRT lines, the marker gene tGFP was part of the target cassette, while DsRed was a split construct completed upon integration of the donor DNA. We transformed BY-2 wild-type suspension cells with agrobacteria containing pRandom and selected 407 events based on imazethapyr and kanamycin resistance (Fig. 1b). To estimate the number of transgene copies in the transformation events, genomic DNA was initially screened by real-time quantitative PCR on both the tGFP and DsRed coding sequence and 35 putative single-copy candidates were selected. To confirm the number of integrated transgene copies in these lines, genomic DNA of all 35 putative single-copy lines was digested with *Nde*I or *Xba*I and analyzed by Southern hybridization with probes binding either the *npt*II coding sequence or the tGFP gene (exemplarily shown in Fig. 1c). Single bands differing in size were detected in 20 of the 35 lines, indicating single-copy integration in these lines at different genome locations (SCR lines; Fig S1). The presence of one or more additional bands in the other 15 lines indicated at least partial multiple copies of the transgenes whether situated at the same site, or elsewhere in the genome

(MCR lines, Fig S1). For the intended analysis of marker gene expression, variability between random and targeted transgene integration events, 12 SCR, and 12 MCR candidate lines were selected based on a reliable growth performance and transgene expression stability. The variance of gene expression in these 24 randomly inserted transformed suspension lines was compared to 15 HRT lines that were generated during a previous study (Schiermeyer et al. 2019) by ZFN-mediated homologous recombination of DNA molecules.

Marker protein fluorescence

Measuring intracellular protein levels and protein distribution within and between cell populations by use of flow cytometry is a very rapid method to detect protein expression on a single-cell level that also allows recordings of multiple proteins in parallel. In our study, each of the cell lines with either targeted (HRT) and random insertions (SCR, MCR) was cultured and used for protoplast preparation. The protoplasts were analyzed by flow cytometry in parallel to determine the amount of DsRed and tGFP marker protein present in the cells. Only protoplasts that were determined to be viable single cells were evaluated for marker protein content; this aim was achieved by gating. This gating showed a high homogeneity of viability among the different HRT lines ($93.7\% \pm 2.9$), while viability in the SCR lines ($68.2\% \pm 17.8$) and the MCR lines ($76.2\% \pm 14.3$) was significantly more heterogeneous (Fig. 2).

As expected, the 15 HRT lines derived from ZFN-targeted gene integration displayed highly homogeneous tGFP levels between the candidate lines and low variability of the geometric mean of fluorescence (FM) between cell lines (FM: 127 ± 17), as indicated by narrow and uniform histograms (Fig. 3a). Among the same HRT lines, the levels of DsRed marker protein were also homogeneous (FM: 4075 ± 2204) but slightly more variable than observed for the tGFP (Fig. 3b). Looking at the 12 SCR lines, the GFP fluorescence is more variable (FM: 124 ± 33) (Fig. 3c) than the GFP fluorescence of the HRT lines, while the DsRed fluorescence is very similar (FM: 4571 ± 2605) to the HRT lines (Fig. 3d). The 12 MCR lines display a distinctly higher variance for both GFP fluorescence (FM: 181 ± 85) and DsRed fluorescence (FM: 6775 ± 4028) (Fig. 3e and f). GFP fluorescence and Dsred fluorescence of all three populations (HRT, SCR, and MCR) are also depicted in direct comparison showing that variation of fluorescence is smallest in HRT lines (Fig. 3g and h).

To test for homogeneity of variance between the three different groups of cell lines, Levenes test ($p \leq 0.05$) based on the assumption that the populations are of equal variance was used. It showed that there was a significant heterogeneity of variance when GFP fluorescence of MCR lines was compared to SCR lines ($p = 0.011$) as well as when it was compared to HRT lines ($p = 0.001$), while there was no significant heterogeneity of variance when GFP fluorescence of SCR lines and HRT lines was compared ($p = 0.32$). For DsRed fluorescence, there was a significant heterogeneity of variance when comparing MCR lines with HRT lines ($p = 0.039$), but no significant heterogeneity of variance between MCR lines and SCR lines ($p = 0.133$) or SCR lines and HRT lines ($p = 0.644$).

To get a better understanding of differences between HRT lines and SCR lines, the standard deviations of DsRed fluorescence in both populations were compared. The average standard deviation of SCR lines (4743 ± 890) is significantly higher ($p = 0.003$) than the average standard deviation of HRT lines (3334 ± 1202), meaning that the homogeneity of cells in each HRT line is significantly higher.

Discussion

The strong interest in industrial biotechnology to generate cell lines with a stable and predictable performance needs laborious processes of screening clones for robust and high production levels. It has been postulated that positional effects have a strong influence on gene expression magnitude and stability both in plants (Day et al. 2000; Matzke and Matzke 1998) as well as in mammalian cell cultures (Perez-Pinera et al. 2012). For plant suspension cultures like tobacco BY-2 cells, no corresponding information has been available yet. Plant suspension cells are an interesting production platform for e.g. pharmaceutical proteins (Moustafa et al. 2016; Schillberg et al. 2013; Topp et al. 2016) and the quick generation of reliable high-producing cell lines would be welcome. If positional effects play a role in protein production in plant suspension cells as well as in whole plants, the use of targeted gene insertion at a specific, pre-evaluated locus would be a highly promising technique. Beside a unique screening platform offering an identical background for e.g. different construct variants, the introduction of transgenes at specific chromosomal locations should eliminate the risk of unpredictable transformation events (Bortesi and Fischer 2015) and thereby has the potential to overcome bureaucratic hurdles in the complex regulatory process and safety analysis of genetically modified plant cells (Hartung and Schiemann 2014; Podevin et al. 2013). Such targeted gene insertion can be achieved by the application of engineered nucleases e.g. the ZFN technology as well as RNA-based methods like the CRISPR/Cas9, which facilitate highly specific gene targeting by triggering the cell's DNA repair machinery. Thus the benefit of reduced screening for high producers and the generation of genetically defined uniform cell lines would be combined with minimal screening efforts (Schiermeyer et al. 2019).
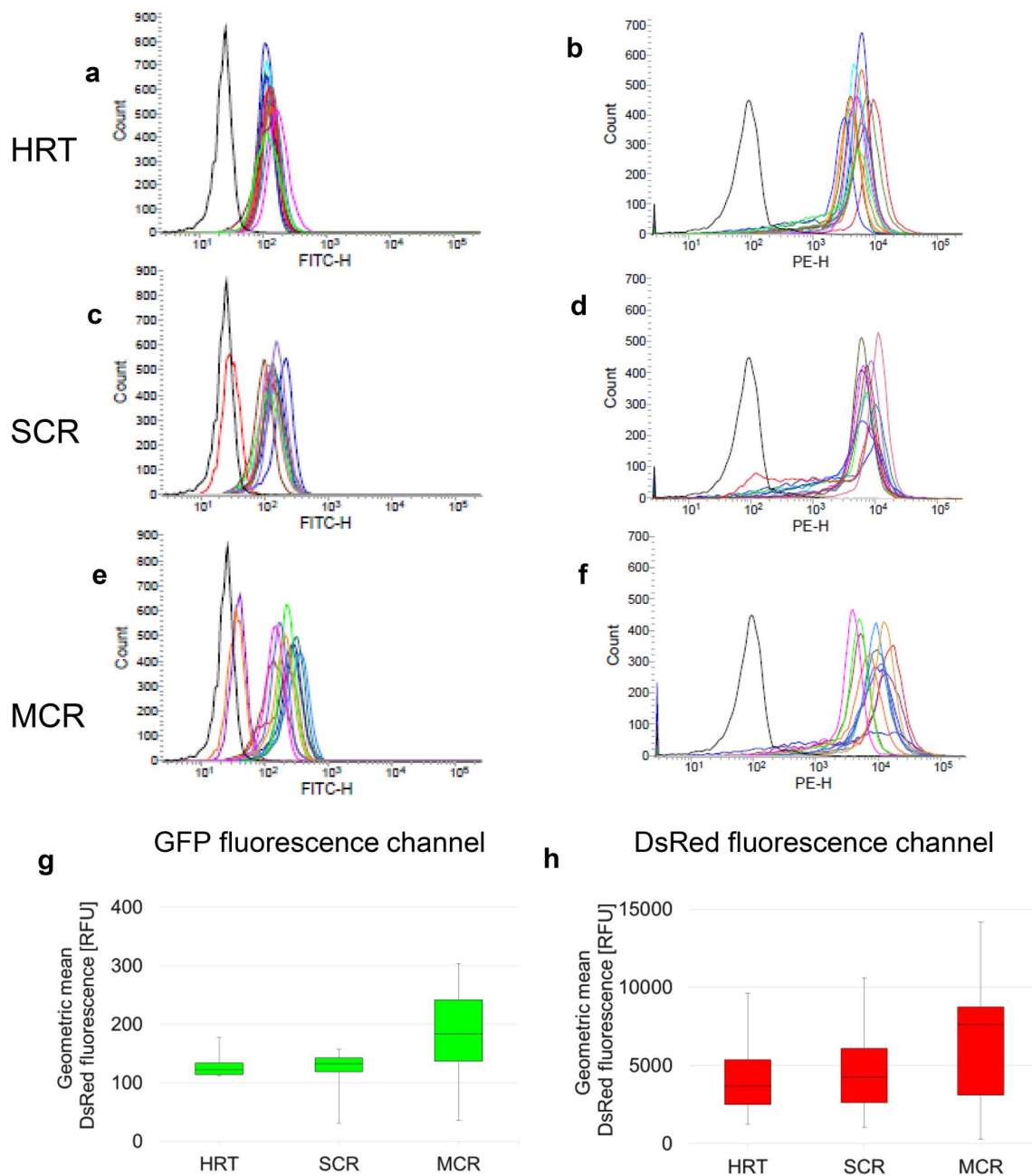
Fig. 3 Analysis of marker protein fluorescence by flow cytometry. To determine the levels of tGFP and DsRed produced, 15 HR-targeted (HRT) and 12 randomly inserted single-copy (SCR) cell lines as well as 12 randomly inserted multiple copy (MCR) cell lines were analyzed by flow cytometry. The fluorescence signal distribution of viable protoplasts for each randomly inserted and HRT cell line is displayed as overlays of histograms for tGFP (**a**, **c**, **e**) and DsRed (**b**, **d**, **f**). Wild-type BY-2 protoplasts were analyzed as internal control and are shown in black. The geometric means of fluorescence values are displayed as boxplot diagrams for tGFP (**g**) and DsRed (**h**) showing the upper 75% and lower 25% percentile values (box), minimum and maximum values (whiskers) and the median/50% percentile values (line)

When analyzing the mean magnitude of fluorescent protein production, no positional effect was observed in this study for single integration events (HRT and SCR), while MCR lines showed enhanced and more variable marker protein levels. This fits with observations by Hobbs et al. (1990) who reported four independent tobacco lines, where single copies of the *uidA* gene had integrated in different chromosome positions but showed similar GUS activity. In the same publication, it is also shown that there are three lines with a double insertion of the transgene that have significantly lower GUS activity. This is also true for some of the multiple insertion lines in this study (e.g. red curve in Fig. 3e and f)

that show a significantly lower tGFP/DsRed fluorescence than the single-copy insertion lines. It is assumed that this effect is due to silencing. Two other reports that support this finding analyzed transgenic Arabidopsis plants with a single copy of the *uidA* gene inserted in different loci (De Buck et al. 2004; Nagaya et al. 2005), but no significant difference in the GUS activity. It should be kept in mind that these analyses were done in plants while we worked with a quickly proliferating plant cell line that is known to be genetically unstable (Kovarik et al. 2012). However, the principle of positional effects not being detectable for the amount of expression of a single-copy transgene seems to hold true.

Another factor that might have an influence on the results when looking for positional effects of single transgene integration is the size of the sample space. While certain studies analyses only up to 21 lines and see no positional effect (De Buck et al. 2004, Hobbs et al. 1990), there is one study (Rosin et al. 2008) that analyzed 277 Arabidopsis lines and noticed approx. 80% to have similar protein production levels, while 20% were subjected to silencing. Silencing was also analyzed in tobacco seedlings that had been transformed with a *gus* reporter gene; here it was noticed that the inserted DNA was methylated in some instances which led to silencing (Day et al. 2000). We have seen this silencing here for the gene coding for tGFP in SCR line 249 (red curve in Fig. 3c and d). This line had been analyzed in a previous experiment 2 months earlier (as had all other lines) and showed at that time a viability of 91% with 82% of the viable protoplasts being red and green, respectively (data not shown). This gives an indication that to truly compare random and targeted insertion in plant cell cultures, it might be necessary to cultivate the cells over a longer period of time and to measure marker protein abundance repeatedly.

The sample space in this study is somewhat limited; the main reason for this is that it was difficult to identify more lines with random insertion that showed a reliable production of the marker proteins. We had identified 20 SCR lines via southern blot analysis (Fig. 1b); however, the number of lines with a reliable cytometry performance was only 12 in the end with most of lines disqualified by missing GFP fluorescence. This could be due to a positional effect as the GFP fluorescence of the HRT lines (the gene in the same locus in all lines) was constant; we started this comparative analysis with 15 lines and did not need to eliminate a single line.

Within the current comparative analysis between protein of marker genes placed in the same genomic locus by targeted integration and single-copy random events, we expected a similar variation of mean fluorescence as observed in mammalian cell cultures (Perez-Pinera et al. 2012) but this is clearly not the case. However, one should keep in mind that marker protein expression will be influenced by the promoter used in the marker gene cassette and the cells in which the study is performed. What we see in the analyzed BY-2 cell lines is that the homogeneity of DsRed fluorescence in each HRT line is significantly higher than in SCR lines; it might be that marker gene production is gradually down regulated in some cells of a SCR line but not in others.

A greater homogeneity is also observed in the viability of analyzed cell lines in HRT lines compared to SCR lines. This is not surprising, as the target cell line TCL448 for HR was carefully selected (Schiermeyer et al. 2019) to guarantee elite characteristics in regard to stable growth performance and GFP production, protoplastation, and cryo-conservation. All HRT lines are essentially the TCL448 and have, therefore, the same elite traits. On the other hand, SCR lines originated from transforming a wild-type culture of BY-2 cells with cells varying in fitness.

The conclusion from the data presented is that while random transformation is a good method when production of proteins is desired, targeted insertion is recommended when a comparison between different constructs is required, as homogeneity and higher reproducibility between experiments of targeted integration into a predesigned landing pad in an elite cell line are clearly superior to random transformation.

References

An GH (1985) High-efficiency transformation of cultured tobacco cells. Plant Physiol 79:568–570

Bevan MW, Flavell RB, Chilton M-D (1983) A chimaeric antibiotic resistance gene as a selectable marker for plant cell transformation. Nature 304:184–187

Bortesi L, Fischer R (2015) The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol Adv 33:41–52

Butaye KMJ, Cammue BPA, Delauré SL, De Bolle MFC (2005) Approaches to minimize variation of transgene expression in plants. Mol Breed 16:79–91

Clendennen SK, Kellogg JA, Phan CB, Mathews HV, Webb NM (2003) Banana and Melon Promoters for Expression of Transgenes in Plants. U.S. Patent Application Publication No. US2003/0182690

Day CD, Lee E, Kobayashi J, Holappa LD, Albert H, Ow DW (2000) Transgene integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced. Genes Dev 14:2869–2880

De Buck S, Windels P, De Loose M, Depicker A (2004) Single-copy T-DNAs integrated at different positions in the Arabidopsis genome display uniform and comparable β-glucuronidase accumulation levels. Cell Mol Life Sci 61:2632–2645

Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM (1982) Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet 1:561–573

Dower WJ, Miller JF, Ragsdale CW (1988) High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res 16:6127–6145

Hartung F, Schiemann J (2014) Precise plant breeding using new genome editing techniques: opportunities, safety and regulation in the EU. Plant J 78:742–752

Hobbs SLA, Kpodar P, Delong CMO (1990) The effect of T-DNA copy number, position and methylation on reporter gene-expression in tobacco transformants. Plant Mol Biol 15:851–864

Kay R, Chan A, Daly M, McPherson J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236:1299–1302

Kirchhoff J, Raven N, Boes A, Roberts JL, Russell S, Treffenfeldt W, Fischer R, Schinkel H, Schiermeyer A, Schillberg S (2012) Monoclonal tobacco cell lines with enhanced recombinant protein yields can be generated from heterogeneous cell suspension cultures by flow sorting. Plant Biotechnol J 10:936–944

Kooter JM, Matzke MA, Meyer P (1999) Listening to the silent genes: transgene silencing, gene regulation and pathogen control. Trends Plant Sci 4:340–347

Kovarik A, Lim K-Y, Soucková-Skalická K, Matyasek R, Leitch AR (2012) A plant culture (BY-2) widely used in molecular and cell studies is genetically unstable and highly heterogeneous. Bot J Linn Soc 170:459–471

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. Methods 25:402–408

Matzke AJ, Matzke MA (1998) Position effects and epigenetic silencing of plant transgenes. Curr Opin Plant Biol 1:142–148

Moustafa K, Makhzoum A, Trémouillaux-Guiller J (2016) Molecular farming on rescue of pharma industry for next generations. Crit Rev Biotechnol 36:840–850

Nagaya S, Kato K, Ninomiya Y, Horie R, Sekine M, Yoshida K, Shinmyo A (2005) Expression of randomly integrated single complete copy transgenes does not vary in *Arabidopsis thaliana*. Plant Cell Physiol 46:438–444

Norris SR, Meyer SE, Callis J (1993) The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. Plant Mol Biol 21:895–906

Peach C, Velten J (1991) Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. Plant Mol Biol 17:49–60

Perez-Pinera P, Ousterout DG, Brown MT, Gersbach CA (2012) Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases. Nucleic Acids Res 40:3741–3752

Podevin N, Davies HV, Hartung F, Nogue F, Casacuberta JM (2013) Site-directed nucleases: a paradigm shift in predictable, knowledge-based plant breeding. Trends Biotechnol 31:375–383

Rajasekaran K, Grula JW, Anderson DM (1996) Selection and characterization of mutant cotton (*Gossypium hirsutum* L.) cell lines resistant to sulfonylurea and imidazolinone herbicides. Plant Sci 119:115–124

Rosin FM, Watanabe N, Cacas JL, Kato N, Arroyo JM, Fang Y, May B, Vaughn M, Simorowski J, Ramu U, McCombie RW, Spector DL, Martienssen RA, Lam E (2008) Genome-wide transposon tagging reveals location-dependent effects on transcription and chromatin organization in Arabidopsis. Plant J 55:514–525

Schiermeyer A, Schneider K, Kirchhoff J, Schmelter T, Koch N, Jiang K, Herwartz D, Blue R, Marri P, Samuel P, Corbin DR, Webb SR, Gonzalez DO, Folkerts O, Fischer R, Schinkel H, Ainley WM, Schillberg S (2019) Targeted addition of large DNA constructs by homology directed repair (HDR) or non-homologous end joining (NHEJ) in engineered tobacco BY-2 cells using designed zinc finger nucleases. Plant Direct. https://doi.org/10.1002/pld3.153

Schillberg S, Raven N, Fischer R, Twyman RM, Schiermeyer A (2013) Molecular farming of pharmaceutical proteins using plant suspension cell and tissue cultures. Curr Pharm Des 19:5531–5542

Schinkel H, Jacobs P, Schillberg S, Wehner M (2008) Infrared picosecond laser for perforation of single plant cells. Biotechnol Bioeng 99:244–248

Schneider K, Schiermeyer A, Dolls A, Koch N, Herwartz D, Kirchhoff J, Fischer R, Russell SM, Cao Z, Corbin DR, Sastry-Dent L, Ainley WM, Webb SR, Schinkel H, Schillberg S (2016) Targeted gene exchange in plant cells mediated by a zinc finger nuclease double cut. Plant Biotechnol J 14:1151–1160

Topp E, Irwin R, McAllister T, Lessard M, Joensuu JJ, Kolotilin I, Conrad U, Stöger E, Mor T, Warzecha H, Hall JC, McLean MD, Cox E, Devriendt B, Potter A, Depicker A, Virdi V, Holbrook L, Doshi K, Dussault M, Friendship R, Yarosh O, Yoo HS, MacDonald J, Menassa R (2016) The case for plant-made veterinary immunotherapeutics. Biotechnol Adv 34:597–604

Wandelt CI, Khan MRI, Craig S, Schroeder HE, Spencer D, Higgins TJV (1992) Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. Plant J 2:181–192