

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Biochemistry -- Faculty Publications

Biochemistry, Department of

---

2016

## An evaluation of new and established methods to determine T-DNA copy number and homozygosity in transgenic plants

Katarzyna Glowacka

*University of Nebraska - Lincoln*, kglowacka2@unl.edu

Johannes Kromdijk

*University of Illinois at Urbana-Champaign*

Lauriebeth Leonelli

*University of California - Berkeley*

Krishna K. Niyogi

*University of California - Berkeley*

Tom E. Clemente

*Center for Plant Science Innovation*, tclemente1@unl.edu

*See next page for additional authors*

Follow this and additional works at: <https://digitalcommons.unl.edu/biochemfacpub>



Part of the [Biochemistry Commons](#), [Biotechnology Commons](#), and the [Other Biochemistry, Biophysics, and Structural Biology Commons](#)

---

Glowacka, Katarzyna; Kromdijk, Johannes; Leonelli, Lauriebeth; Niyogi, Krishna K.; Clemente, Tom E.; and Long, Stephen P., "An evaluation of new and established methods to determine T-DNA copy number and homozygosity in transgenic plants" (2016). *Biochemistry -- Faculty Publications*. 424.

<https://digitalcommons.unl.edu/biochemfacpub/424>

This Article is brought to you for free and open access by the Biochemistry, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Biochemistry -- Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

---

**Authors**

Katarzyna Glowacka, Johannes Kromdijk, Lauriebeth Leonelli, Krishna K. Niyogi, Tom E. Clemente, and Stephen P. Long

## Original Article

## An evaluation of new and established methods to determine T-DNA copy number and homozygosity in transgenic plants

Katarzyna Głowacka<sup>\*1,2</sup>, Johannes Kromdijk<sup>\*1</sup>, Lauriebeth Leonelli<sup>3</sup>, Krishna K. Niyogi<sup>3,4</sup>, Tom E. Clemente<sup>5</sup> & Stephen P. Long<sup>1</sup><sup>1</sup>Carl R. Woese Institute for Genomic Biology, University of Illinois, 1206 W Gregory Drive, Urbana, IL 61801, USA, <sup>2</sup>Institute of Plant Genetics, Polish Academy of Sciences, ul. Strzeszyńska 34, 60-479 Poznań, Poland, <sup>3</sup>Howard Hughes Medical Institute, Department of Plant and Microbial Biology, 111 Koshland Hall, University of California Berkeley, Berkeley, CA 94720-3102, USA, <sup>4</sup>Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA and <sup>5</sup>Center for Plant Science Innovation, E324 Beadle Center, 1901 Vine Street, Lincoln, NE 68588, USA

## ABSTRACT

**Stable transformation of plants is a powerful tool for hypothesis testing. A rapid and reliable evaluation method of the transgenic allele for copy number and homozygosity is vital in analysing these transformations. Here the suitability of Southern blot analysis, thermal asymmetric interlaced (TAIL-)PCR, quantitative (q)PCR and digital droplet (dd)PCR to estimate T-DNA copy number, locus complexity and homozygosity were compared in transgenic tobacco. Southern blot analysis and ddPCR on three generations of transgenic offspring with contrasting zygosity and copy number were entirely consistent, whereas TAIL-PCR often underestimated copy number. qPCR deviated considerably from the Southern blot results and had lower precision and higher variability than ddPCR. Comparison of segregation analyses and ddPCR of T<sub>1</sub> progeny from 26 T<sub>0</sub> plants showed that at least 19% of the lines carried multiple T-DNA insertions per locus, which can lead to unstable transgene expression. Segregation analyses failed to detect these multiple copies, presumably because of their close linkage. This shows the importance of routine T-DNA copy number estimation. Based on our results, ddPCR is the most suitable method, because it is as reliable as Southern blot analysis yet much faster. A protocol for this application of ddPCR to large plant genomes is provided.**

**Key-words:** ddPCR; digital droplet PCR; qPCR; segregation analysis; selectable marker; Southern blot; TAIL-PCR; transformation.

## INTRODUCTION

The use of stable transformation of plants to test a range of hypotheses has been increasing rapidly and has been suggested a key element in addressing future security of food supply as well as in adapting to global change (Khush 2005; Hibberd *et al.* 2008; Long *et al.* 2015). In parallel, there is considerable interest in introducing new pathways and modifying existing pathways in plants to produce new or improved bioproducts (Clemente & Cahoon 2009; Rogers & Oldroyd 2014). Rapid and reliable evaluation of transgenic allele(s)

for copy number and homozygosity is a vital step in utilizing these transformation events, so that homozygous lines with stable transgene expression are obtained for accurate testing. Knowledge of the transgenic locus structure and zygosity status is particularly important in plants, including all crops, which have a life cycle of months and sometimes years. Here, decreasing the number of generations between the initial transformation and identification of homozygous lines with stable expression has greatest value.

Southern blot analysis (Southern, 1975), quantitative polymerase chain reaction (qPCR; Higuchi *et al.* 1992), thermal asymmetric interlaced-PCR (TAIL-PCR; Liu *et al.* 1995) and most recently digital droplet PCR (ddPCR; Hindson *et al.* 2011) have all been used to provide information about the integration status of a transgenic allele(s) in genomes (Tingay *et al.* 1997; Fu *et al.* 1999; Campbell *et al.* 2000; Ingham *et al.* 2001; German *et al.* 2003; Pillai *et al.* 2008; Mieog *et al.* 2013; Yi *et al.* 2013; Larkan *et al.* 2016; Wang *et al.* 2015). However, these methods use contrasting principles for characterization of a transgenic allele and differ vastly in precision, reproducibility and potential to scale up. Southern blot analysis has been regarded as the most unambiguous method for molecular characterization of transgenic alleles for both estimation of copy number and loci complexity in transgenic plants. The disadvantage is that the process is more labour intensive, and less well suited for automation, relative to PCR-based platforms. Here the precision and speed of different PCR methods are compared to Southern blot analysis.

Typically, a large number of independent transformations are generated to compensate for insertions in undesirable genomic loci; for example insertion into a native open reading frame, or in a region not conducive to stable expression levels. The resulting collection of primary transformants can then be screened to prioritize lead events that possess targeted transgene expression levels, coupled with simple integration and low copy transgenic allele(s). Importantly, an ideal genotyping platform should possess the ability to facilitate identification of homozygous lineages in early generations of selection.

In Southern blot analysis (Southern, 1975) genomic DNA is digested, separated on a gel, blotted onto a membrane and then detected with a radioactive, fluorescent or chemiluminescent labelled probe sequence to visualize the transgenic allele(s) complexity of integration. The intensity of

Correspondence: S. P. Long. +1 217 244 0881; e-mail: slong@illinois.edu  
\*Joint First Authors.

the hybridization signal in reconstruction lanes, whereby known quantities of the target sequences are spiked into native DNA, is correlative with zygosity at each allele (Dai *et al.* 2001). In qPCR (Higuchi *et al.* 1992), template concentration is estimated based on the fluorescence trace from a dye or probe in the sample which is directly proportional to the template amplification. Copy number, as well as zygosity, of the transgenic allele(s) can be derived from qPCR analysis, but logarithmic PCR template quantification may limit ability to identify small copy number differences (Bubner & Baldwin 2004). Additionally, direct coupling between the PCR amplification and quantification makes qPCR very sensitive to PCR efficiency (Bustin *et al.* 2009). TAIL-PCR (Liu *et al.* 1995) has also been used to establish the number of insertions in transgenic events by means of amplifying the flanking sequence of the transgenic allele(s), which can be exploited to map the allele if a genome sequence is available (e.g. Hanhineva & Kärenlampi 2007). Recently, a ddPCR method has been developed (Hindson *et al.* 2011) which can also be used to detect absolute DNA copy concentration at high accuracy. The detection principle is based on a fluorescent dye or probe, similar to qPCR. However in ddPCR, the PCR reaction is rendered digital by randomization of the sample DNA over a set of subsamples, in which the DNA dilution is chosen to obtain both positive and negative subsamples for template presence. A Poisson probability distribution can then be used to derive the template concentration (Hindson *et al.* 2011). Decoupling of amplification and quantification should make ddPCR results relatively insensitive to PCR efficiency and the linearity of the quantification scale should allow detection of small copy differences (Hindson *et al.* 2013, Bharuthram *et al.* 2014). While segregation analysis of transgene presence in subsequent generations can provide information about the mode of inheritance, and a means to identify homozygous lines for down-stream phenotyping, this approach requires several generations of inbred plants to be analysed and associated costs for time, space and labour can be substantial. Additionally, although *Agrobacterium tumefaciens*-mediated transformation typically results in a lower T-DNA copy number than direct DNA transfer methods such as electroporation or particle bombardment, multiple T-DNAs integrated at the same locus can still be found (Jorgensen *et al.* 1987; Kohli *et al.* 1999; Gelvin 2003). Such multiple inserts at a single locus would not be detected by segregation analysis. Multiple T-DNA copies can increase the likelihood of silencing effects, in particular when tandem T-DNA insertions are arranged in inverted repeats (Jorgensen *et al.* 1987; Stam *et al.* 1997). This makes detecting these multiple inserts of particular importance in selecting for stable expression.

The aim of this study was to assess both the importance of estimating copy number and to assess which method is most suited to estimating copy numbers and establishing homozygosity in transgenic plants. T-DNA copy number in the same transgenic plants was evaluated with Southern blot, qPCR, TAIL-PCR and ddPCR analyses. Subsequently we assessed reproducibility of parallel qPCR and ddPCR

analyses, using both methods with 12 different primer combinations on the same DNA sample. After comparing data output, duration of protocol, reproducibility and precision, the ddPCR method appeared most suitable to routinely assess T-DNA copy numbers and homozygosity. An example protocol optimized for use in higher plants with large genomes is provided in the Supporting Information.

## MATERIAL AND METHODS

### Plant material and transformation

*Nicotiana tabacum* cv. 'Petite Havana' was transformed using the *Agrobacterium*-mediated leaf disc protocol according to Clemente (2006), a protocol based on the seminal communication by Horsch *et al.* (1985), using two different plasmids. The first plasmid (PsbS) contained the coding sequence of the photosystem II subunit S gene from *Nicotiana benthamiana* (*NbPsbS*). Overexpression of PsbS results in a proportional increase in non-photochemical quenching (NPQ) of chlorophyll fluorescence in transformed plants (Li *et al.* 2002). Therefore chlorophyll fluorescence analysis could be used to identify transformed individuals in segregation analysis. The second plasmid (VPZ) contained coding sequences of three genes from *Arabidopsis thaliana*: violaxanthin de-epoxidase (*AtVDE*), *AtPsbS* and zeaxanthin epoxidase (*AtZEP*). T-DNA maps are provided in Supporting Information (Fig. S1A–B). The presence of three genes in transformants with construct 2 allowed us to verify reproducibility of the qPCR and ddPCR method with multiple primer sets to derive the copy number of the transgenic cassette. Additionally, both constructs contained the *bar* gene encoding resistance for bialaphos (Thompson *et al.* 1987). Multiple confirmed independent T<sub>0</sub> transformants were generated for each construct (11 and 15 for PsbS and VPZ, respectively).

### DNA extraction for PCR methods

Young leaf tissue was collected from greenhouse-grown plants to determine T-DNA insert copy number, snap frozen in liquid nitrogen and stored in  $-20^{\circ}\text{C}$ . DNA was extracted by the CTAB method modified from Kabelka *et al.* (2002).

### Southern blot

Southern hybridization was carried out on a subset of eight plants derived from one VPZ transformation event (VPZ-23) representing T<sub>0</sub>, T<sub>1</sub> and T<sub>2</sub> generations. These were used to compare Southern blot analysis with the three PCR based methods. Genomic DNA was extracted using a modified version of the protocol by Dellaporta *et al.* (1983). Fifteen micrograms of genomic DNA was digested overnight with *Bam*HI (R3136, New England Biolabs, Ipswich, MA, USA) and separated on a 0.8% agarose gel at 25 V overnight and alkali blotted onto a Zeta-probe GT genomic tested blotting membrane according to the manufacturer's specifications (BioRad, Hercules, CA, USA). Probe DNA was obtained by double digestion of the VPZ plasmid using *Xho*I (R0146, New England Biolabs, Ipswich, MA, USA) and

*Xba*I (R0145, New England Biolabs, Ipswich, MA, USA), yielding a 594 bp fragment, which was radiolabelled with  $\alpha$ -<sup>32</sup>P-labelled dCTP by means of the Prime-It II random primer labeling kit according to the manufacturer's protocol (Stratagene, La Jolla, CA, USA). Pre-hybridization and hybridization were carried out at 65 °C in the presence of high salt buffer following the manufacturer specifications. Membranes were analysed by autoradiography (X-Omat AR5 film, Eastman Kodak, Rochester, NY, USA).

### Thermal asymmetric interlaced PCR (TAIL-PCR)

DNA for thermal asymmetric interlaced (TAIL-) PCR was extracted as described above. TAIL-PCR conditions were the same as described in Liu *et al.* (1995), except that primary and secondary reaction volumes were 25  $\mu$ L instead of 20  $\mu$ L and tertiary reaction volumes 50  $\mu$ L instead of 100  $\mu$ L. Additionally, in all three reactions the Phusion high fidelity PCR master-mix with high fidelity buffer (M0531S, New England Biolabs, Ipswich, MA, USA) was used together with appropriate primer set with 2  $\mu$ M of arbitrary degenerate primer (AD1, AD2 or AD3) and 200 nmole of T-DNA specific primer (TR1, TR2 or TR3; for primer sequences see Table S1). The amplicons of all three reactions were separated on 1% agarose gels (General purpose agarose GP2, BE-A125, MidSci, Valley Park, MO, USA) with Tris-Borate-EDTA buffer and visualized with ethidium bromide.

### Quantitative PCR (qPCR)

DNA obtained as described above was digested overnight with *Hind*III (R3104, New England Biolabs, Ipswich, MA, USA), followed by purification with cleaning and concentrating columns (D4014, Zymo Research, Irvine, CA, USA). Reactions were prepared using 9  $\mu$ L of digested genomic DNA (20–25 ng  $\mu$ L<sup>-1</sup>), 200 nmole of forward and reverse primersets AtPsbS\_3, AtPsbS\_4, AtVDE\_1, AtVDE\_4, AtZEP\_1 or AtZEP\_4 for T-DNA amplicons and NtActin\_1 and NtTubulin\_1 for reference genes (for primer sequences, see Supporting Information Table S1) and 10  $\mu$ L of SsoAdvanced Universal SYBER Green Supermix (172-5270; BioRad, Hercules, CA, USA). The reaction efficiency for each primer set was estimated on a series of DNA dilutions. All reactions were run on CFX connect Real-Time PCR Detection System (1855201, BioRad, Singapore) using the following program: 3 min 95 °C, 40 times (10 s 95 °C; 30 s 60 °C), followed by melting curve generated from 65 °C to 90 °C. Four technical replicates were used for both T-DNA and the reference sequence. Raw data was processed using BioRad CFX Manager 3.1 and T-DNA copy number estimates made using the  $\Delta\Delta$ Ct method (Livak & Schmittgen, 2001) and corresponding standard errors were computed as described in Hoebeek *et al.* (2007).

### Digital droplet PCR (ddPCR)

Digested DNA was obtained as described for qPCR, excepting that in addition to *Hind*III. *Sac*I (R3156, New England Biolabs,

Ipswich, MA, USA) was also used for the digestion of DNA derived from PsbS transformants. The full step-by-step protocol for ddPCR is described in the Supporting Information (Appendix 1). Briefly, the ddPCR reactions were prepared using 11.5  $\mu$ L of digested genomic DNA (20–25 ng  $\mu$ L<sup>-1</sup>), 100–150 nmole of forward and reverse primers (for primer sequences, see Supporting Information Table S1), 12.5  $\mu$ L of a commercial reaction mix including an intercalating fluorescent dye, polymerase, Mg<sup>2+</sup> and dNTPs (2x QX200 ddPCR EvaGreen Supermix, 186-4034, BioRad, Hercules, CA, USA) and MilliQ autoclaved water (dMQ H<sub>2</sub>O) to a total reaction volume of 25  $\mu$ L. Droplets were generated using droplet generator cartridges, gaskets and cartridge holder (186-4007 and DG8, 186-3051, Bio-Rad, Hercules, CA, USA) in a droplet generator (QX200, 186-4002, Bio-Rad, Hercules, CA, USA). According to the manufacturer's instructions, 20  $\mu$ L of ddPCR reaction mix and 70  $\mu$ L of droplet generation oil (EvaGreen 186-4005, Bio-Rad, Hercules, CA, USA) were loaded into cartridges. The resulting 40  $\mu$ L sample of generated droplets was dispensed to one well of a 96-well PCR semi-skirted plate (951020362, Eppendorf, Enfield, CT, USA) and sealed with foil (Pierceable foil heat seal, 181-4040 and Px1 PCR plate sealer, 181-4000, Bio-Rad, Hercules, CA, USA). The droplet mix was cycled through a PCR program using a deep-well thermal cycler (C1000 Touch, 185-1196, Bio-Rad, Hercules, CA, USA), followed immediately by analysis in a droplet reader (QX200, 186-4003, Bio-Rad, Hercules, CA, USA).

Fluorescence reads per individual droplet from each well were analysed with the manufacturer-provided software (Quanta Soft version 1.7, 1864011, Bio-Rad, Hercules, CA, USA). The droplet population in each well was divided into template negative or positive. The distribution of positive and negative droplets is a function of the starting PCR template concentration according to the Poisson probability distribution, which allows the absolute concentration of PCR-template to be derived from the number of negative droplets according to Eqn 1:

$$\text{PCR Concentration – template} = \frac{-\ln(N_{\text{neg}}/N)}{V} \quad (1)$$

where  $N_{\text{neg}}$  is the number of negative droplets;  $N$  is the total number of droplets and  $V$  is volume of a single droplet which is equal to 0.85 nL. For the calculation of copy number the ratio between the concentration of the T-DNA and a reference sequence of known copy number was used. Two primer sets were used to amplify reference sequences in either  $\alpha$ -tubulin (two copies) or actin (four copies). Primer sequences were designed to amplify regions in contig c61851 and c50972, respectively from *N. tabacum* (Methylation Filtered Genome TGI: v.1 Contigs; accessed August 4, 2014). The performance of each primer set was empirically verified prior to use. T-DNA copy number and corresponding standard errors were calculated based on the ratio of T-DNA versus reference sequence according to Hedges *et al.* (1999). Two technical replicates were used for both T-DNA and reference gene ddPCR reactions. In all comparative analyses, the same primers were used for qPCR and ddPCR.

## Segregation analysis according to glufosinate ammonium resistance

T<sub>1</sub> progeny of 11 P<sub>sb</sub>S lines and 15 VPZ lines were used in segregation studies of resistance to glufosinate ammonium. Seeds collected from T<sub>0</sub> plants were germinated in round 20-cm-diameter pots (ITML Inc, Brantford, ON, Canada) containing a peat-, bark- and perlite-based growing medium (Metro-Mix 900, Sun Gro Horticulture, Agawam, MA, USA). Five days after germination, 24 randomly chosen seedlings of each line were transferred to twelve 4 cm × 3 cm × 5.7 cm cells of planting trays (715364C, T.O. Plastics, MN, USA) placed on flats (710251C, T.O. Plastics, MN, USA) filled with a growing medium with a reduced perlite content (Metro-Mix 510, Sun Gro Horticulture, Agawam, MA, USA). Seedlings were grown in a controlled environment walk-in growing chamber with 12 h day (23 °C)/12 h night (18 °C) cycle under 150 μmol quanta m<sup>-2</sup> s<sup>-1</sup>. Ten days after transplanting, seedlings were sprayed with 1.5 g L<sup>-1</sup> glufosinate ammonium solution [0.8% (v/v) Ignite herbicide, cas77182-82-7, Bayer CropScience, Research Triangle Park, NC, USA]. Presence of necrotic and chlorotic leaf tissue was assessed 10 days after exposure and used to score plants as either susceptible or resistant. Additionally, 72 plants of each of five T<sub>1</sub> lines were grown in the field, and mature leaves of each plant were painted with the same solution and then assessed for damage 10 days later. In both cases, the observed susceptible plant counts were compared against the expected counts based on 1:3 segregation using a Chi<sup>2</sup> test (α = 0.1).

## Segregation analysis using NPQ level in P<sub>sb</sub>S-transformed plants

The abundance of P<sub>sb</sub>S protein is tightly linked to the maximum level of rapidly reversible NPQ of Photosystem II (NPQ; Li *et al.* 2002). This property allows use of NPQ level as an independent marker for progeny segregation. Ten-day-old seedlings of T<sub>1</sub> progeny of 11 P<sub>sb</sub>S lines were germinated and grown, as described in the preceding section, and used in modulated chlorophyll fluorescence imaging to determine NPQ (CF Imager, Technologica, Colchester, UK). Seedlings were dark adapted for 20 min and then the dark-adapted maximal fluorescence (F<sub>m</sub>) was imaged with a 800 ms saturating flash (6000 μmol quanta m<sup>-2</sup> s<sup>-1</sup>) from blue LEDs (λ<sub>max</sub> = 470 nm). Immediately following the dark-adapted reading, blue LEDs were powered to provide 1000 μmol quanta m<sup>-2</sup> s<sup>-1</sup> on the seedlings. After 10 min a second saturating flash was applied to determine the maximum fluorescence under illuminated conditions (F<sub>m</sub>'). Average NPQ per seedling was then calculated from these measurements according to Eqn 2, assuming the Stern–Volmer quenching model (Maxwell and Johnson 2000):

$$\text{NPQ} = \text{F}_m/\text{F}_m' - 1. \quad (2)$$

To separate transgenic T<sub>1</sub> progeny into those carrying the P<sub>sb</sub>S construct or segregated back to wild-type (WT), the mean and standard deviation of WT seedling NPQ values were used to identify a threshold level of NPQ, above which seedlings were classified as transgenic (α ≤ 0.05). The resulting transgenic

and WT counts for each T<sub>1</sub> progeny were subsequently used to estimate the number of independently inherited loci, by comparing against the expected counts based on 1:3 segregation using a Chi<sup>2</sup> test (α = 0.1).

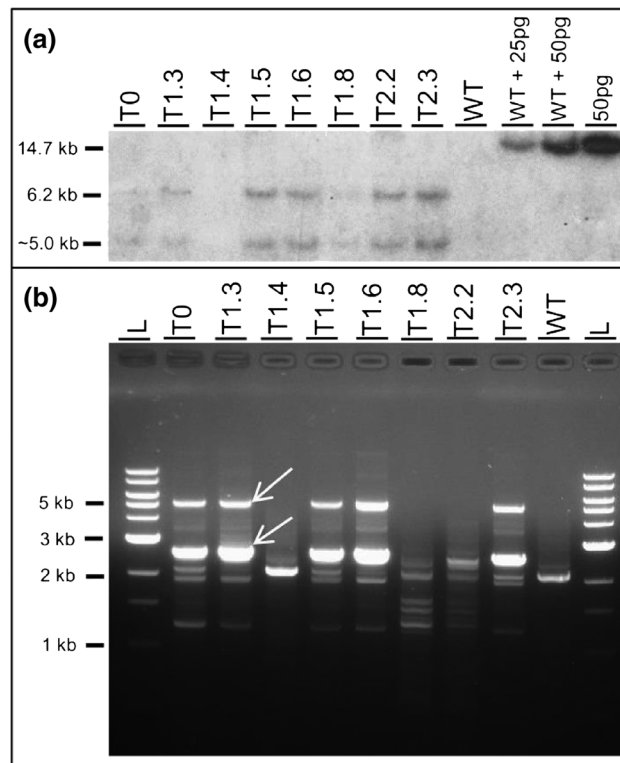
## RESULTS

### Comparison between southern blot analysis, TAIL-PCR, qPCR and ddPCR

Southern blot analysis on selected VPZ-23 plants showed two bands of approximately 5.0 and 6.2 kb in the VPZ-23 T<sub>0</sub> plant (lane 1), suggesting two T-DNA copies present in the primary transformant (Fig. 1a; Table 1). Self-pollination resulted in T<sub>1</sub> progeny with zero (T<sub>1.4</sub>), two T-DNA copies (T<sub>1.3</sub> and T<sub>1.8</sub>) or four copies (T<sub>1.5</sub> and T<sub>1.6</sub>). TAIL-PCR was performed on the same plants using AD1, AD2 and AD3 primers with increasing degeneracy (64-, 128- and 256-fold, respectively) (Table 1, Fig. 1b and Fig. S2). Results with AD1 showed only one unique band for plants T<sub>0</sub>, T<sub>1.3</sub>, T<sub>1.5</sub>, T<sub>1.6</sub> and T<sub>2.3</sub>, whereas the second band expected based on Southern blot results was missing and DNA from plants T<sub>1.8</sub> and T<sub>2.2</sub> failed to generate any unique band. Increasing the degeneracy of the primer to 256-fold resulted in one unique band for plants T<sub>1.8</sub> and T<sub>2.2</sub>, and two unique bands for plants T<sub>0</sub>, T<sub>1.3</sub>, T<sub>1.5</sub>, T<sub>1.6</sub> and T<sub>2.3</sub>. The corresponding T-DNA copy number from qPCR showed no copies in WT and T<sub>1.4</sub>, and clear differences between plants T<sub>0</sub>, T<sub>1.3</sub> and T<sub>1.8</sub> (2.17–2.37), and plants T<sub>1.5</sub>, T<sub>1.6</sub>, T<sub>2.2</sub> and T<sub>2.3</sub> (4.4–5.37) (Table 1). However, in general the T-DNA copy numbers were higher than Southern blot estimates by 0.17 to 1.37, with standard errors ranging from 0.06 (T<sub>0</sub>) to 0.62 (plant T<sub>1.8</sub>). Finally, T-DNA copy number estimates from ddPCR were very closely approximated to the corresponding Southern blot results with 1.83–1.90 for plants with two copies and 3.74–4.24 for plants with four copies. In contrast to qPCR results, no systematic over-estimation was present in the ddPCR results relative to the Southern blot analysis. Standard errors were generally lower for ddPCR than qPCR, ranging between 0.04 (plants T<sub>1.5</sub> and T<sub>1.8</sub>) to 0.28 (plant T<sub>1.6</sub>) even though only two technical replicates were used compared to four for qPCR, thus more strongly reflecting the likelihood of two copies per locus as visualized by the Southern hybridization (Fig. 1; Table 1).

### Reproducibility of ddPCR and qPCR estimates

Because qPCR and ddPCR appeared to have greater accuracy than TAIL-PCR, an additional side-by-side comparison of these two methods was performed to evaluate reproducibility. Using DNA from a plant transformed with the VPZ construct, two different primer sets targeted to contrasting regions per gene (*AtVDE*, *AtPsbS* and *AtZEP*) were used, together with primers targeting regions in two different reference sequences (α-tubulin and actin). These reference sequences have different copy numbers in the tobacco genome, two versus four, respectively. The qPCR efficiency for these eight primer sets ranged from 97 to 102%. The resulting 12 combinations of transgene and reference primer sets and the corresponding estimated



**Figure 1.** (a) Southern blot (b) TAIL-PCR analyses for T<sub>0</sub> plant VPZ-23, five segregating T<sub>1</sub> plants, two homozygous T<sub>2</sub> plants and wild type control (WT). The final three lanes show 25 and 50 pg digested VPZ plasmid DNA with 10 μg of digested WT DNA and 50 pg of VPZ plasmid without WT DNA. Arrows in panel b indicate the bands that were absent in WT and show a size shift between reaction 2 and 3 in the TAIL-PCR and scored in Table 1. TAIL-PCR was performed with AD3 and T-DNA specific primers RB3.

copy number using qPCR and ddPCR are shown in Table 2. T-DNA copy number estimates varied between 1.49 and 5.38 for qPCR but only 1.82 and 2.26 for ddPCR. Estimates using  $\alpha$ -tubulin as a reference sequence appeared consistently higher when determined by qPCR, whereas no obvious bias could be detected for either actin or  $\alpha$ -tubulin in

ddPCR. As in the previous comparison, standard errors for the estimated T-DNA copy numbers were generally higher for qPCR (0.05–0.49) compared to ddPCR (0.02–0.12), even though fewer technical replicates were used for ddPCR; two versus four. Because all analyses should arrive at the same answer, we computed the variance across the 12

**Table 1.** Comparison of the T-DNA copy numbers estimated by Southern blot, qPCR, TAIL-PCR and ddPCR for eight plants representing three generation (T<sub>0</sub>, T<sub>1</sub> and T<sub>2</sub>) and corresponding control wild type (WT) of *N. tabaccum*. For TAIL-PCR the T-DNA copy number was assessed with the use of three different reaction sets with primers differing in degeneracy: AD1, AD2 and AD3, with 64-, 128- and 256-fold degeneracy, respectively. For qPCR and ddPCR the T-DNA copy number was derived from the estimated sample concentration of T-DNA relative to actin ( $\pm$  se;  $n = 4$  for qPCR;  $n = 2$  for ddPCR).

Plant ID	T-DNA copy number					
	Southern blot	qPCR	TAIL-PCR			ddPCR
			AD1	AD2	AD3	
VPZ-23 T <sub>0</sub>	2	2.37 ± 0.06	1	1	2	1.83 ± 0.11
VPZ-23 T <sub>1.3</sub>	2	2.17 ± 0.18	1	1	2	1.90 ± 0.20
VPZ-23 T <sub>1.4</sub>	0	0 ± 0.00	0	0	0	0 ± 0.00
VPZ-23 T <sub>1.5</sub>	4	4.83 ± 0.19	1	1	2	3.74 ± 0.04
VPZ-23 T <sub>1.6</sub>	4	5.37 ± 0.15	1	1	2	4.21 ± 0.28
VPZ-23 T <sub>1.8</sub>	2	2.29 ± 0.62	0	1	1	1.87 ± 0.04
VPZ-23 T <sub>2.2</sub>	4	4.40 ± 0.12	0	0	1	4.24 ± 0.11
VPZ-23 T <sub>2.3</sub>	4	4.72 ± 0.16	1	1	2	4.18 ± 0.16
WT	0	0 ± 0.00	0	0	0	0 ± 0.00

**Table 2.** Comparison of T-DNA copy number estimated using ddPCR or qPCR estimations of PCR template concentration of six different primer sets targeting T-DNA regions and either actin or  $\alpha$ -tubulin as a reference gene. Analysis was performed on digested DNA of VPZ-46 T<sub>1.5</sub> plant carrying a T-DNA with three genes (*AtVDE*, *AtPsb*s and *AtZEP*). ( $\pm$  se;  $n = 4$  for qPCR;  $n = 2$  for ddPCR).

Primer set name	qPCR		ddPCR	
	T-DNA copy number		T-DNA copy number	
	actin	$\alpha$ -tubulin	actin	$\alpha$ -tubulin
AtPsbS_3	1.49 $\pm$ 0.11	2.85 $\pm$ 0.25	2.01 $\pm$ 0.05	1.82 $\pm$ 0.03
AtPsbS_4	1.86 $\pm$ 0.49	2.29 $\pm$ 0.06	2.19 $\pm$ 0.06	1.98 $\pm$ 0.03
AtVDE_1	2.08 $\pm$ 0.26	3.19 $\pm$ 0.13	2.26 $\pm$ 0.12	2.05 $\pm$ 0.06
AtVDE_4	3.51 $\pm$ 0.20	5.38 $\pm$ 0.10	2.06 $\pm$ 0.05	1.87 $\pm$ 0.02
AtZEP_1	2.01 $\pm$ 0.18	3.08 $\pm$ 0.09	2.16 $\pm$ 0.08	1.96 $\pm$ 0.04
AtZEP_4	1.92 $\pm$ 0.09	2.94 $\pm$ 0.05	2.16 $\pm$ 0.11	1.96 $\pm$ 0.06
Variance per method	1.10		0.02	

combinations per method in order to evaluate reproducibility. The variance across the qPCR determinations was substantially higher than across the ddPCR estimates, 1.10 versus 0.02.

### What is the most suitable method to assess T-DNA copy number and homozygosity?

Table 3 shows that all four methods can generate T-DNA copy number estimates in T<sub>0</sub> transformants, although the precision is low for qPCR and TAIL-PCR. Additionally, Southern blot analysis, qPCR and ddPCR were effective in estimating loci number and identifying homozygous offspring in analysis of T<sub>1</sub> progeny. Southern blots were taken as the benchmark for precision in determining copy number. Although both qPCR and ddPCR matched Southern blot determinations, qPCR gave a greater variance and also tended to over-estimate copy number. TAIL-PCR clearly under-estimated copy number, and this was improved little by use of the most degenerate primer, AD3. The comparison between ddPCR and qPCR using a set of 12 different combinations of T-DNA and reference sequence primer sets, showed good reproducibility in the ddPCR estimates, whereas qPCR results were much more variable, leading to contrasting results between the different primer set combinations (Table 2). A key consideration for

all methods is the time that they require. TAIL-PCR requires three consecutive PCR programmes, which all have to be verified using gel electrophoresis, and together add up to approximately 12 h. Both qPCR and ddPCR protocols made use of pre-digested DNA, which was done overnight; hence, the total protocol as described here would take up to one day. The estimate for Southern blot analysis is considerably longer, accounting for the exposure time of the x-ray film to the labelled membrane. In addition, all PCR methods can be easily automated using 96-well plates or greater, allowing many samples to be run in this time. The largest Southern blot gel would typically be 18 sample tracks; this could be multiplied but automation is considerably more difficult (Table 3).

### Screening for homozygous plants in T<sub>1</sub> progeny

Because ddPCR results were highly reproducible and consistent with Southern blot analysis, we further evaluated ddPCR as a means to rapidly identify homozygosity in T<sub>1</sub> progeny. Because T<sub>0</sub> plants are hemizygous, cross-comparison between T<sub>0</sub> and T<sub>1</sub> T-DNA copy numbers allows easy identification of homozygous plants in T<sub>1</sub>, from which the WT segregants had already been eliminated by glufosinate ammonium treatment. In *N. tabacum*, which is tetraploid but a functional diploid, homozygous plants in

**Table 3.** Comparison of time required, precision, reproducibility and form of the output data of four different methods used in T-DNA copy estimation.

Feature	Method			
	Southern blot	qPCR	TAIL-PCR	ddPCR
Time required <sup>b</sup>	4–7 days	1.5 day <sup>a</sup>	2 days	1.5 day <sup>a</sup>
Precision	High	Low	Low	High
Reproducibility	n.e. <sup>c</sup>	Low	n.e. <sup>c</sup>	High
Type of output data				
T-DNA copy number	+	+	-/+	+
T-DNA locus number	+	+	-	+
Homozygosity of T <sub>1</sub> plants	-/+	+	-	+

<sup>a</sup>It is possible to make these protocols faster by shortening the pre-digestion time.

<sup>b</sup>Estimate does not include the additional time to generate T<sub>1</sub> progeny, which will be very species dependent.

<sup>c</sup>Not estimated.



the  $T_1$  offspring should show a doubling of the T-DNA copy number relative to the corresponding  $T_0$  plant. Using this

**Table 4.** Results of screening for copy number to identify homozygous plants in  $T_1$  segregating progenies with the use of ddPCR. Seedlings of  $T_1$  lines were sprayed with glufosinate ammonium solution before single plants were transferred to pots and grown for tissue collection for ddPCR analyses. Bold numbers indicate homozygous plants identified by a duplication of the copy number in the corresponding  $T_0$  plant ( $\pm$  se;  $n = 2$ ).

Generation	Transformation event			
	PsbS-43	PsbS-46	VPZ-29	VPZ-34
$T_0$	1.09 $\pm$ 0.13	0.98 $\pm$ 0.18	0.97 $\pm$ 0.09	0.97 $\pm$ 0.14
$T_1$	<b>2.10 <math>\pm</math> 0.22</b>	<b>2.04 <math>\pm</math> 0.16</b>	<b>1.96 <math>\pm</math> 0.16</b>	0.94 $\pm$ 0.21
	<b>2.13 <math>\pm</math> 0.20</b>	1.08 $\pm$ 0.11	0.93 $\pm$ 0.15	<b>1.93 <math>\pm</math> 0.06</b>
	<b>2.05 <math>\pm</math> 0.04</b>	1.01 $\pm$ 0.28	0.98 $\pm$ 0.16	<b>2.08 <math>\pm</math> 0.06</b>
	1.12 $\pm$ 0.25	<b>2.08 <math>\pm</math> 0.10</b>	1.04 $\pm$ 0.20	0.98 $\pm$ 0.09
	<b>2.02 <math>\pm</math> 0.16</b>	<b>2.09 <math>\pm</math> 0.08</b>	1.03 $\pm$ 0.27	<b>1.80 <math>\pm</math> 0.21</b>
	<b>2.04 <math>\pm</math> 0.05</b>	<b>2.04 <math>\pm</math> 0.05</b>	0.95 $\pm$ 0.08	<b>1.87 <math>\pm</math> 0.11</b>
	<b>2.26 <math>\pm</math> 0.04</b>	<b>2.10 <math>\pm</math> 0.08</b>	1.03 $\pm$ 0.10	<b>1.80 <math>\pm</math> 0.31</b>
	<b>2.10 <math>\pm</math> 0.28</b>	1.08 $\pm$ 0.26	1.02 $\pm$ 0.03	<b>2.02 <math>\pm</math> 0.20</b>

**Table 5.** T-DNA copy ( $\pm$  se;  $n = 2$ ) estimated by digital droplet PCR (ddPCR) and estimated number of T-DNA loci based on segregation analyses ( $\chi^2$ -test for 1:3 segregation,  $\alpha = 0.1$ ). Segregation was scored based on either resistance to glufosinate ammonium (GA) in 17-day-old seedlings or level of non-photochemical quenching (NPQ) in 7-day-old seedlings. NPQ values were induced by 10 min exposure to PFD of 1000  $\mu\text{mol quanta m}^{-2} \text{s}^{-1}$ .

ID of $T_0$ plant	ddPCR No. of T-DNA copy	Segregation of resistance to GA leaf paint in $T_1$ generation			Segregation based on NPQ in Fig. 2b		
		No. of susceptible plants (%)	No. of resistant plants (%)	No. of T-DNA loci in line	No. of plants with WT NPQ (%)	No. of plants with NPQ higher than WT (%)	No. of T-DNA loci in line
PsbS-23	0.99 $\pm$ 0.17	6 (27)	16 (73)	1	2 (11)	16 (89)	1
PsbS-25	0.98 $\pm$ 0.06	8 (33)	16 (67)	1	2 (12)	15 (88)	1
PsbS-27	0.97 $\pm$ 0.04	5 (23)	17 (77)	1	7 (41)	10 (59)	1
PsbS-28	1.04 $\pm$ 0.10	5 (21)	19 (79)	1	3 (17)	15 (83)	1
PsbS-34	0.87 $\pm$ 0.22	4 (17)	20 (83)	1	1 (5)	18 (95)	>1
PsbS-43	1.09 $\pm$ 0.13	5 (21)	19 (79)	1	4 (22)	14 (78)	1
PsbS-47	0.98 $\pm$ 0.18	3 (12.5)	21 (87.5)	1	4 (21)	15 (79)	1
PsbS-49	1.99 $\pm$ 0.21	0 (0)	23 (100)	>1	1 (6)	17 (94)	>1
PsbS-50	1.97 $\pm$ 0.10	3 (12.5)	21 (87.5)	1	4 (22)	14 (78)	1
PsbS-2	3.94 $\pm$ 0.21	0 (0)	24 (100)	>1	0 (0)	15 (100)	>1
		1 1 <sup>a</sup>	71 (99) <sup>a</sup>	>1 <sup>a</sup>			
PsbS-32	3.89 $\pm$ 0.05	6 (26)	17 (74)	1	4 (29)	10 (71)	1
VPZ-14	1.03 $\pm$ 0.38	5 (21)	19 (79)	1	na	na	na
		19 (26) <sup>a</sup>	53 (74) <sup>a</sup>	1 <sup>a</sup>			
VPZ-29	0.97 $\pm$ 0.09	6 (26)	17 (74)	1	na	na	na
VPZ-31	1.07 $\pm$ 0.09	5 (21)	19 (79)	1	na	na	na
VPZ-34	0.97 $\pm$ 0.14	8 (33)	16 (67)	1	na	na	na
VPZ-38	1.02 $\pm$ 0.05	2 (9)	21 (91)	>1	na	na	na
VPZ-54	0.97 $\pm$ 0.07	9 (37.5)	15 (62.5)	1	na	na	na
VPZ-13	2.12 $\pm$ 0.08	5 (23)	17 (77)	1	na	na	na
		20 (28) <sup>a</sup>	51 (72) <sup>a</sup>	1 <sup>a</sup>			
VPZ-23	1.83 $\pm$ 0.11	4 (17)	20 (83)	1	na	na	na
		18 (25) <sup>a</sup>	54 (75) <sup>a</sup>	1 <sup>a</sup>			
VPZ-28	1.91 $\pm$ 0.16	2 (8)	22 (92)	>1	na	na	na
VPZ-35	2.94 $\pm$ 0.13	1 (4)	23 (96)	>1	na	na	na
VPZ-50	2.86 $\pm$ 0.05	1 (4)	23 (96)	>1	na	na	na
VPZ-52	2.88 $\pm$ 0.17	2 (8)	22 (92)	>1	na	na	na
VPZ-51	4.83 $\pm$ 0.14	4 (17)	20 (83)	1	na	na	na
VPZ-33	11.96 $\pm$ 0.13	15 (62.5)	9 (37.5)	na <sup>b</sup>	na	na	na
VPZ-36	12.09 $\pm$ 0.08	0 (0)	24 (100)	>1	na	na	na

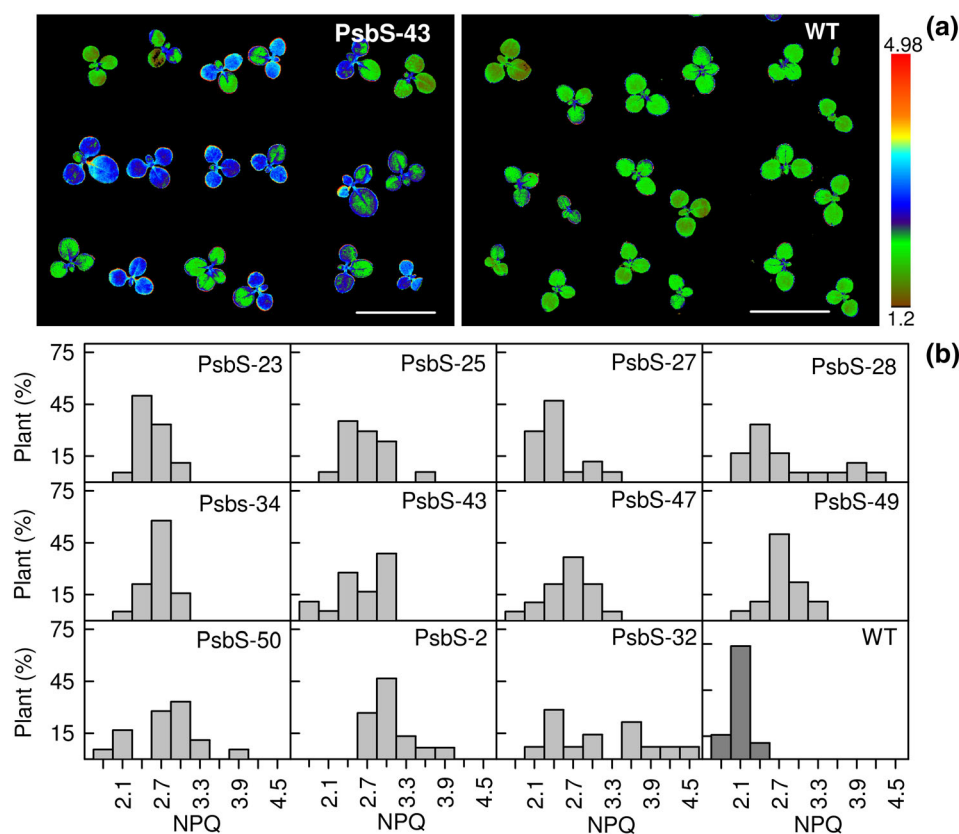
<sup>a</sup>Data collected of resistance to leaf-paint of GA solution in 72 field-grown plants.

<sup>b</sup>Not estimated because of non-Mendelian segregation. na – not applicable.

criterion, Table 4 shows that ddPCR provides an unambiguous identification of the homozygotes in the 32  $T_1$  progeny.

### Comparison of T-DNA copy numbers and numbers of T-DNA loci

Assuming Mendelian inheritance, segregation analysis is another means to obtain homozygous lines; however, if multiple insertions occur at the same locus or are otherwise linked, this will fail to show true copy number. For 26 independent transformations, 17 of these showed approximately 1:3 segregation of glufosinate ammonium resistance, indicating a single T-DNA insertion locus. However, ddPCR showed that five of these 17 carried more than one T-DNA copy, and in one case 5. This suggests a high frequency of multiple inserts at a single locus that would not be detected by segregation analysis (Table 5). Chlorophyll fluorescence imaging of NPQ was used as an easily screenable phenotype in assessing the PsbS transformant segregation, of which an example is shown in Fig. 2a. PsbS transformants were easily distinguishable from WT based on increased NPQ (Fig. 2b), and segregation



**Figure 2.** The segregation of nonphotochemical quenching (NPQ) in 10-day-old seedlings transformed by *NbPsbS* plants of *N. tabacum*. (a) Imaged NPQ for PsbS-43 T<sub>1</sub> (segregating T<sub>1</sub> progeny of T<sub>0</sub> plant carrying one T-DNA copy) and wild type control (WT); (b) distribution of NPQ in 11 T<sub>1</sub> segregating populations and WT. Presented values of NPQ were recorded after 10 min of induction at 1000  $\mu\text{mol quanta m}^{-2} \text{s}^{-1}$ . Bar on panel (a) represents 2.5 cm.

analysis gave very similar locus number results compared to glufosinate ammonium resistance (Table 5).

## DISCUSSION

### Evaluation of different methods to establish T-DNA copy number

Although Southern blots have historically been used to estimate transgenic allele loci and copy number, the method typically takes at least a week to generate results, relatively few samples can be processed in parallel and require specialized radiation laboratories. The PCR-based methods presented in the current manuscript took typically less than two days, can be easily optimized to quantify final copy number in a matter of hours and can easily be automated. Therefore, PCR-based methods are attractive because of reduced time to results and general ease of use. Using Southern blots as a benchmark, we show that T-DNA loci determination and estimate of copy numbers predicted with the ddPCR method (Table 1) were most accurate across a subset of three generations of transgenic plants, with contrasting zygosity and copy number and generally showed the lowest variance in estimates. Estimates from qPCR were more variable, and consistently overestimated copy numbers relative to the Southern blot estimates, while

TAIL-PCR underestimated copy number (Table 1). Similar problems have been shown for copy number estimation based on inverse PCR, which is also based on amplification of flanking sites (Does *et al.* 1991). In TAIL-PCR, closely spaced T-DNA may decrease the statistical probability of a suitable binding site for the degenerate primer between subsequent T-DNA insertions. If this was the case, using a higher fold of degeneracy should improve the number of amplified flanking sites. In six of the assessed samples, using a more degenerate primer indeed did improve the number of bands identified (see AD1 to AD3, Table 1), but in two samples the number of bands was still lower than expected based on Southern blot results. Therefore, TAIL-PCR may be used reliably to detect at least one T-DNA copy, but the frequent underestimation relative to ddPCR and Southern blot estimates is difficult to troubleshoot when *a priori* knowledge of the correct copy number is not already available.

In a separate comparison, ddPCR was also shown to be more reproducible than qPCR in a parallel analysis of 12 different combinations of target and reference primer sets on DNA from a single plant (Table 2). Results for ddPCR consistently indicated two T-DNA copies, independent of the primer combination used. However, qPCR estimates showed a range of contrasting T-DNA copy numbers, varying between 1.49 and 5.38, such that interpretation of the T-DNA copy number was

unreliable and tended to over-estimation. These results can be explained by a number of advantages favouring ddPCR over qPCR for the purpose of T-DNA copy number estimation. Contrary to qPCR, in ddPCR the level of fluorescence is not directly coupled to the PCR target quantification, allowing PCR efficiency requirements to be less strict. As a result ddPCR is more robust against factors interfering with PCR target amplification and no standard curve is required. The linearity of the ddPCR technique also allows precise detection of small fold changes in copy concentration, whereas the logarithmic detection scale of qPCR limits the capabilities to accurately detect small differences. A number of medical studies have recently compared qPCR and ddPCR for absolute quantification of human microRNAs and copy number variation. These studies also found ddPCR to have far less variability and increased accuracy compared to qPCR (Hindson *et al.* 2013; Bharuthram *et al.* 2014).

### T-DNA copy number and number of independent inherited loci

Because  $T_0$  plants carrying a single T-DNA locus are most easily progressed to a fully homozygous generation, in most cases these primary transformants are preferred for further study. Table 4 shows that homozygous individuals in four lines with a single T-DNA can be easily distinguished. The ratio of homozygous versus hemizygous  $T_1$  progeny in Table 4 is slightly higher than the expected 1:2. Because the  $T_1$  offspring was first treated with glufosinate ammonium solution to eliminate the WT fraction, this offset may be explained by a possible *bar* gene dosage effect between the hemizygous and homozygous lines, possibly favouring the latter to be identified for further screening.

The consistency of predictions between independent data for the well-established selectable marker glufosinate ammonium resistance and NPQ levels in the PsbS transformants (Table 5) shows that co-transformation of the PsbS gene can be used as a visual marker, which avoids issues with toxicity and resistance outcrossing often associated with traditional markers (Miki & McHugh, 2004). In 5 out of 26  $T_0$  plants, copy numbers assessed by ddPCR were higher than the loci number estimated by segregation analyses of glufosinate ammonium resistance and NPQ. These  $T_0$  plants were all estimated to have a single independent T-DNA locus based on segregation analyses (Table 5), whereas ddPCR results showed up to five T-DNA copies (VPZ-51) were integrated. These findings may be explained by the presence of silent or incomplete T-DNA copies, which would be detected by ddPCR but not affect the inheritance pattern of the selectable marker. Alternatively, these results suggest a relatively high occurrence of multiple T-DNA copies per locus (at least 19% in our study), which is consistent with earlier reports (Jorgensen *et al.* 1987; Kohli *et al.* 1999; Gelvin 2003) and shows the importance of routinely checking the T-DNA copy number, even if segregation analysis suggests the occurrence of only one T-DNA containing locus. Because multiple T-DNA copies at the same locus can increase the occurrence of silencing, these findings exemplify the importance to allow precise and routine

selection of single copy  $T_0$  transformants for further analysis, which cannot be substituted by segregation analysis. We used two different constructs, with contrasting T-DNA lengths (4.5 kb and 10.8 kb for PsbS and VPZ, respectively), but the length of the T-DNA seemed to make little difference in the occurrence of multiple T-DNA insertions.

The iterative cycle of modification, testing and evaluation which is essential in genetic engineering approaches relies on high-throughput generation and verification of transgenic plants. Our results show ddPCR to be accurate, precise and fast in determining T-DNA copy numbers for screening the high numbers of transgenic plants being produced today across a wide range of research goals. The protocol (Supporting Information; Appendix 1) can also easily be modified to allow automated high-throughput screening of DNA samples. Our study shows serious limitations in applying qPCR and TAIL-PCR in estimating copy number. The results also show a high frequency of genetically linked insertions which could not be detected by segregation analyses, emphasizing the importance of routinely checking T-DNA copy numbers in generation of transgenic plants. ddPCR provides a high-throughput means to achieve this.

### ACKNOWLEDGEMENTS

This research was supported by the Bill and Melinda Gates Foundation (OPP1060461) titled 'RIPE – Realizing Increased Photosynthetic Efficiency for Sustainable Increases in Crop Yield'. K.K.N. is an investigator of the Howard Hughes Medical Institute and the Gordon and Betty Moore Foundation (through Grant GBMF3070). We thank Prof S. Whitney for seeds of *N. tabacum* cv 'Petite Havana'. Finally, we also thank David Drag, Ben Harbaugh, Steven Huber Jr, Brittiani Batts and Lynnica Massenburg for help with collecting seeds and tissue samples.

### REFERENCES

- Bharuthram A., Paximadis M., Picton A. & Tiemessen C. (2014) Comparison of a quantitative Real-Time PCR assay and droplet digital PCR for copy number analysis of the CCL4L genes. *Infection, Genetics and Evolution* **25**, 28–35.
- Bubner B. & Baldwin I.T. (2004) Use of real-time PCR for determining copy number and zygosity in transgenic plants. *Plant Cell Reports* **23**, 263–271.
- Bustin S.A., Benes V., Garson J.A., Hellems J., Huggett J., Kubista M. & Wittwer C.T. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* **55**, 611–622.
- Campbell B.T., Baenziger P.S., Mitra A., Sato S. & Clemente T. (2000) Inheritance of multiple transgenes in wheat. *Crop Science* **40**, 1133–1141.
- Clemente T. (2006) *Nicotiana* (Nicotiana tabacum, *Nicotiana benthamiana*). In *Agrobacterium Protocols* (ed Wang K.), pp. 143–154. Humana Press Inc., Totowa.
- Clemente T.E. & Cahoon E.B. (2009) Soybean oil: genetic approaches for modification of functionality and total content. *Plant Physiology* **151**, 1030–1040.
- Dai S., Zheng P., Marmey P., Zhang S., Tian W., Chen S. & Fauquet C. (2001) Comparative analysis of transgenic rice plants obtained by Agrobacterium-mediated transformation and particle bombardment. *Molecular Breeding* **7**, 25–33.
- Dellaporta S.L., Wood J. & Hicks J.B. (1983) A plant DNA miniprep: version II. *Plant Molecular Biology Reporter* **1**, 19–21.
- Does M.P., Dekker B.M.M., de Groot M.J.A. & Offringa R. (1991) A quick method to estimate the T-DNA copy number in transgenic plants at an early stage after transformation, using inverse PCR. *Plant Molecular Biology* **17**, 151–153.

- Fu X., Duc L.T., Fontana S., Bong B.B., Tinjuangjun P., Sudhakar D. & Kohli A. (1999) Linear transgene constructs lacking vector backbone sequences generate low-copy-number transgenic plants with simple integration patterns. *Transgenic Research* **9**, 11–19.
- Gelvin S.B. (2003) Agrobacterium-mediated plant transformation: the biology behind the “gene-jockeying” tool. *Microbiology and Molecular Biology Reviews* **67**, 16–37.
- German M.A., Kandel-Kfir M., Swarzburg D., Matsevitz T. & Granot D. (2003) A rapid method for the analysis of zygosity in transgenic plants. *Plant Science* **164**, 183–187.
- Hanhineva K.J. & Kärenlampi S.O. (2007) Production of transgenic strawberries by temporary immersion bioreactor system and verification by TAIL-PCR. *BMC Biotechnology* **7**, 11.
- Hedges L.V., Gurevitch J. & Curtis P.S. (1999) The meta-analysis of response ratios in experimental ecology. *Ecology* **80**, 1150–1156.
- Hibberd J.M., Sheehy J.E. & Langdale J.A. (2008) Using C4 photosynthesis to increase yield of rice – rationale and feasibility. *Current Opinion in Plant Biology* **11**, 228–231.
- Higuchi R., Dollinger G., Walsh S.P. & Griffith R. (1992) Simultaneous amplification and detection of specific dna sequences. *Biotechnology* **10**, 413–417.
- Hindson B.J., Ness K.D., Masquelier D.A., Belgrader P., Heredia N.J., Makarewicz A.J. & Colston B.W. (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Analytical Chemistry* **83**, 8604–8610.
- Hindson C.M., Chevillet J.R., Briggs H.A., Gallichotte E.N., Ruf I.K., Hindson B.J. & Tewari M. (2013) Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nature Methods* **10**, 1003–1005.
- Hoebeek J., Speleman F. & Vandesompele J. (2007) Real-time quantitative PCR as an alternative to southern blot or fluorescence in situ hybridization for detection of gene copy number changes. In *Protocols for Nucleic Acid Analysis by Nonradioactive Probes* (eds Hilario E. & Mackay J.), pp. 205–226. Humana Press Inc., Totowa.
- Horsch R.B., Fry J.E., Hoffmann N.L., Eichholtz D., Rogers S.G. & Fraley R.T. (1985) A simple and general method for transferring genes into plants. *Science* **227**, 1229–1231.
- Ingham D.J., Beer S., Money S. & Hansen G. (2001) Quantitative real-time PCR assay for determining transgene copy number in transformed plants. *BioTechniques* **31**, 132–140.
- Jorgensen R., Snyder C. & Jones J.D.G. (1987) T-DNA is organized predominantly in inverted repeat structures in plants transformed with *Agrobacterium tumefaciens* C58 derivatives. *Molecular Genomics and Genetics* **207**, 471–477.
- Kabelka E., Franchino B. & Francis D. (2002) Two loci from *Lycopersicon hirsutum* LA407 confer resistance to strains of *Clavibacter michiganensis* subsp. *Michiganensis*. *Phytopathology* **9**, 504–510.
- Khush G.S. (2005) What will it take to feed 5.0 billion rice consumers? *Plant Molecular Biology* **59**, 1–6.
- Kohli A., Gahakwa D., Vain P., Laurie D.A. & Christou P. (1999) Transgene expression in rice engineered through particle bombardment: molecular factors controlling stable expression and transgene silencing. *Planta* **208**, 88–97.
- Larkan N.J., Ma L. & Borhan M.H. (2016) The *Brassica napus* receptor-like protein RLM2 is encoded by a second allele of the *LepR3/Rlm2* blackleg resistance locus. *Plant Biotechnology Journal* **13**, 983–992.
- Li X.P., Gilmore A.M. & Niyogi K.K. (2002) Molecular and global time-resolved analysis of a *psbS* gene dosage effect on pH- and xanthophyll cycle-dependent nonphotochemical quenching in photosystem II. *Journal of Biological Chemistry* **277**, 33590–33597.
- Liu Y.-G., Mitsukawa N., Oosumi T. & Whittier R.F. (1995) Efficient isolation and mapping of Arabidopsis T-DNA insert junctions by thermal asymmetric interlaced PCR. *The Plant Journal* **8**, 457–463.
- Livak K.J. & Schmittgen T.D. (2001) Analysis of relative gene expression data using realtime quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *METHODS* **25**, 402–408.
- Long S.P., Marshall-Colon A. & Zhu X.-G. (2015) Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **161**, 56–66.
- Maxwell K. & Johnson G.N. (2000) Chlorophyll fluorescence – a practical guide. *Journal of Experimental Botany* **51**, 659–668.
- Mieog J.C., Howitt C.A. & Ral J.-P. (2013) Fast-tracking development of homozygous transgenic cereal lines using a simple and highly flexible real-time PCR assay. *BMC Plant Biology* **13**, 71.
- Miki B. & McHugh S. (2004) Selectable marker genes in transgenic plants: applications, alternatives and biosafety. *Journal of Biotechnology* **107**, 193–232.
- Pillai M.M., Venkataraman G.M., Kosak S. & Torok-Storb B. (2008) Integration site analysis in transgenic mice by thermal asymmetric interlaced (TAIL)-PCR: segregating multiple-integrant founder lines and determining zygosity. *Transgenic Research* **17**, 749–754.
- Rogers C. & Oldroyd G.E. (2014) Synthetic biology approaches to engineering the nitrogen symbiosis in cereals. *Journal of Experimental Botany* **65**, 1939–1946.
- Southern E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–517.
- Stam M., De Bruin R., Kenter S., Van der Hoorn R.A.L., Van Blokland R., Mol J. M.N. & Kooter J.M. (1997) Post-transcriptional silencing of chalcone synthase in *Petunia* by inverted transgene repeats. *The Plant Journal* **12**, 63–82.
- Thompson A.C., Nicollier G.F. & Pope D.F. (1987) Indolealkylamines of *Desmanthus illinoensis* and their growth inhibition activity. *Journal of Agricultural and Food Chemistry* **35**, 361–365.
- Tingay S., McElroy D., Kalla R., Fieg S., Wang M., Thornton S. & Brettell R. (1997) *Agrobacterium tumefaciens*-mediated barley transformation. *The Plant Journal* **11**, 1369–1376.
- Wang X., Jiang D. & Yang D. (2015) Fast-tracking determination of homozygous transgenic lines and transgene stacking using a reliable quantitative real-time PCR assay. *Applied Biochemistry and Biotechnology* **175**, 996–1006.
- Yi D., Cui L., Wang L., Liu Y., Zhuang M., Zhang Y. & Yang L. (2013) Pyramiding of Bt *cry1Ia8* and *cry1Ba3* genes into cabbage (*Brassica oleracea* L. var. *capitata*) confers effective control against diamondback moth. *Plant Cell Tissue and Organ Culture* **115**, 419–428.

Received 21 October 2015; accepted for publication 9 December 2015

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

**Table S1.** Sequence of primers used in ddPCR, qPCR and TAIL-PCR to assess copy number of T-DNA inserts.

**Figure S1.** Maps of T-DNA of PsbS and VPZ constructs.

**Figure S2.** Agarose gel of primary, secondary and tertiary TAIL-PCR products.

**Appendix 1.** ddPCR protocol.docx