Auxin independent1 (*Axi1*) as an endogenous gene for copy number determination in transgenic tobacco

Simara da Silva Lopes[1], Jéssica Fabiane da Veiga Costa[2], Beatriz de Almeida Barros[3], Andrea Almeida Carneiro[3] and Sylvia Morais de Sousa[1, 2, 3*]

Abstract: *Determining the copy number is important because it can greatly influence the expression level and genetic stability of transgenes. This study aimed to establish a methodology that can estimate the transgene copy number in tobacco (Nicotiana tabacum) using the Auxin independent1 (Axi1) gene as an endogenous control. Tobacco c.v. Petit Havana plants were transformed via Agrobacterium tumefaciens with a pMCG1005 vector that contained Bar as a selective gene. Oligonucleotide efficiency was determined and a qPCR using the $2^{-\Delta Ct}$ and $2^{-\Delta\Delta Ct}$ methods was performed. Bar was the target gene and Axi1 was the endogenous control in five transgenic tobacco events. The results showed that the Axi1 gene was donated by the maternal parent N. sylvestris when interspecific hybridization occurred between N. sylvestris and N. tomentosiformis. The copy number results agreed with the segregation ratios for the Bar gene in T1 plants, which confirmed that Axi1 is a single copy gene that can be used as an endogenous control.*

Keywords: *Bar, Nicotiana tabacum, real-time PCR, segregation, transgene*

INTRODUCTION

Genetic engineering incorporates new genes into transgenic plants from completely unrelated species, which can be transmitted to their descendants (Low et al. 2018, Nazir et al. 2019). In 1983, the first transgenic tobacco (*Nicotiana tabacum*) plants containing a gene coding for resistance against the antibiotic kanamycin were generated (Fraley et al. 1983). Since then, several proteins have been expressed in transgenic tobacco and other plants (Jube and Borthakur 2007, Burnett and Burnett 2019, Moon et al. 2020). Tobacco is a highly productive, non-food crop that is cultivated in about 120 countries throughout the world. Furthermore, it is considered to be a model plant for scientific research because of its well-established regeneration and transformation methodologies, and the existence of a sequenced genome (Edwards et al. 2017).

The genetic transformation of model plants is key to validating gene function. Transgenic plants can be used to characterize single genes and reconstruct gene networks that control plant biochemistry, physiology, and morphology under different conditions (Kochetov and Shumny 2017). Molecular characterization is necessary after transformation because copy number and the insertion site can influence gene stability and the expression levels of the

*Corresponding author:
E-mail: sylvia.sousa@embrapa.br
ORCID: 0000-0002-8030-1385

[1] Universidade Federal de São João del-Rei, 36.301-160, São João Del-Rei, MG, Brazil
[2] Centro Universitário de Sete Lagoas, 35.701-242, Sete Lagoas, MG, Brazil
[3] Embrapa Milho e Sorgo, 35.701-970, Sete Lagoas, MG, Brazil

transgene (Gadaleta et al. 2011). Transgenic (T_0) plants are hemizygous for the gene of interest and if they inherit a dominant trait and follow Mendelian inheritance, then they will produce homozygous and hemizygous plants after segregation. Out of all the generated plants, only one quarter of the T1 generation are homozygous plants that can be used to improve crops. Therefore, it is important to find simple, reliable, and high-throughput techniques that can detect the zygosity of transgenes (Passricha et al. 2016). Southern blotting was used for a long time to determine copy numbers, but it is a laborious and time-consuming method. It requires considerable amounts of DNA from samples and hazardous radioisotopes in some cases. Real-time PCR was identified as a fast, sensitive, and accurate technique that could be used to determine copy number insertion. However, in order to ensure that this method produces accurate results, it is essential to have a single copy gene as a reference, such as *Alcohol dehydrogenase 1* (*Adh1*) in maize (Ingham et al. 2001) and a gene encoding the I/Y protein of a mobile element (*hmg l/y*) in *Brassica napus* (Masek et al. 2000, Weng et al. 2005).

Not many single genes in tobacco have been characterized as being potential single-copy reference genes. The *Auxin independent1* (*Axi1*) gene was first used to determine the copy number in transgenic plants transformed with the *p3* gene (Potato Virus A - PVA), an aphid-transmitted potyvirus (Nováková et al. 2005) by Šubr et al. (2006). However, they were unable to prove that *Axi1* was a single copy-gene. The *Axi1* gene was isolated from a DNA tagged plant mutant collection that aimed to isolate genes involved in auxin action. When *Axi1* expression is interrupted, the protoplasts gain the ability to grow in culture in the absence or at high concentrations of auxin, which is uncommon in wild type protoplasts (Walden et al. 1994).

Tobacco is an allotetraploid and its large genome (4.5 Gb) contains a large number of replicates (> 70%). The species *N. tabacum* (2n = 4x = 48) evolved from the interspecific hybridization of the ancestors *Nicotiana sylvestris* (2n = 24, maternal donor) and *Nicotiana tomentosiformis* (2n = 24, paternal donor) about 400,000-800,000 years ago (Clarkson et al. 2017). The results from this study showed that *Axi1* is a single copy gene and that it came from the maternal donor *N. sylvestris*. We also established a methodology to estimate transgene copy number in transgenic tobacco using real-time PCR and the $2^{-\Delta\Delta Ct}$ method.

MATERIAL AND METHODS

Cloning the *Axi1 gene*

Oligonucleotides were designed and used to clone a 727 bp fragment from the *Axi1* gene (GenBank: X80301) (*Axi1*CloneF: 5'-AGATGCAGTTGTTGCAGCTC-3' and *Axi1*CloneR: 5'-TCAGATGCAAGGCAACAAAG-3'). The PCR reaction solution contained 40 ng tobacco genomic DNA (20 ng mL^{-1}), 0.5 mM MgCl$_2$, 0.5 mM of each primer, and 1× Promega GoTaq® Colorless Master Mix in a final volume of 20 mL. The reaction solutions were heated to 94 °C for 2 min, followed by 35 cycles of 94 °C for 20 s, 61 °C for 20 s, and 72 °C for 1 min. The final step was 72 °C for 5 min. The PCR product was applied to 1% agarose gel and visualized with GelRed (Biotium, Hayward, CA, USA). The band from the fragment of interest was isolated from the gel and purified using a QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) and then cloned into the pGEM®-T Easy Vector System (Promega Corporation, Madison, WI, USA). *Escherichia coli* DH5-a cells were transformed using the freeze-thaw method and the plasmid DNA from the recombinant clones was extracted using a Wizard® Plus SV Minipreps DNA Purification System kit (Promega Corporation, Madison, WI, USA). The plasmid DNA was quantified using a Nanodrop1000 spectrophotometer (Thermo Fisher Scientific, Foster City, CA, USA). All procedures were performed according to manufacturer's instructions.

Determination of the *Axi1 gene* copy number in the tobacco genome

The copy number of the *Axi1* gene in the tobacco genome was determined using a standard four-point curve obtained from a 1:10 factor serial dilution of plasmid DNA containing 100 000, 10 000, 1000, and 100 copies of the *Axi1* gene. The copy number for the plasmid (starting amount) was calculated using the formula:

$$Copy\ Number\ \mu L^{-1} = \frac{molecular\ mass}{mass\ (g)} \times 6.022 \times 10^{23}$$

The CT values were plotted against the log of the copy number and the standard curve was generated from an adjusted

linear regression of the plotted points. The PCR amplification efficiency (E) was calculated from the slope coefficient of the adjusted line using the equation in Rasmussen (2001):

$$10^{-1}/_{slope\,-1}$$

The copy number of the gene *Axi1* in the tobacco genome was determined by taking 87 ng (10 000 copies of the genome – content 2C = 8.7 pg) from three independent genomic DNA samples. The results were then interpolated using the standard curve.

The PCR reaction sample contained plasmid DNA, and there were three biological replicates and three technical replicates with 1:10 factor serial dilution of plasmid DNA (100 000, 10 000, 1000, and 100 copies of the *Axi1* gene). The sample also contained 87 ng genomic DNA, 1× Fast SYBR® Green Master Mix (Applied Biosystems, Thermo Fisher Scientific), 1 mM of each primer (*Axi1*Forward: 5´-GCCGTCCTTGTAGTTCCAAA-3´ and *Axi1*Reverse:5´-AGCGGTCGACATCAAAAATC-3´) (Šubr et al. 2006) in a final volume of 10 mL. The *Axi1* primers were located on exons five (forward) and six (reverse), and they amplified an 85 bp section of DNA. A 7500 Fast Real-Time PCR System (Applied Biosystems, Thermo Fisher Scientific) was used for the PCR and the amplification conditions were 95 °C for 20 s, followed by 40 cycles of 95 °C for 3 s, and 60 °C for 30 s, which were the manufacturer's recommendations.

Confirmation of the *Axi1* parental donor

Genomic DNA from *N. tabacum* leaves and its parents, *N. sylvestris* and *N. tomentosiformis*, was extracted using the Saghai-Maroof et al. (1984) method. A 975 bp fragment was amplified from 40 ng DNA (20 ng mL^{-1}) in a mixture that contained 0.6 mM MgCl$_2$ and 1 mM *Axi1*F2 (5´- CCAAAATCATCATCATCAACCA-3') and *Axi1*R2 (5'-TCTGTTGATTCAAGCCTCCA-3'), 5% DMSO, 1.25× Promega GoTaq® Colorless Master Mix in a final volume of 20 mL. The cycle conditions were 94 °C for 2 min, followed by a 35 cycles of 94 °C for 20 s, 59 °C for 20 s, 72 °C for 1 min, and a final extension of 72 °C for 5 min.

Tobacco transformation with the pMCG1005 vector

Agrobacterium tumefaciens strain EHA101 harboring a pMCG1005 binary vector (McGinnis et al. 2005) that expressed the *Bar* gene driven by the 4× 35S promoter and the nopaline synthase gene terminator was used to transform axenic *Nicotiana tabacum* cv. Petit Havana plants following the protocol developed by Horsch et al. (1985). Tobacco leaf disks were placed in a Petri dish that was lined with tissue. Then 20 mL *Agrobacterium tumefaciens* culture (grown at 28 °C, 250 rpm) was added and the samples were left to culture overnight. After culturing, the leaf disks were removed and any excess bacterial culture was blotted from the tissue using sterile filter paper. The leaf disks were placed onto infection medium (4.3 g L^{-1} MS salts mixture, 3% sucrose, 0.8% agar, 0.4 mg L^{-1} thiamine-HCl, 100 mg L^{-1} myo-inositol, 5 mM kinetin, and 13.5 mM 2,4-dichlorophenoxyacetic acid) and incubated for 48 hours at 26 °C in a lit growth chamber. Unless otherwise stated, all of the chemicals were obtained from the Sigma Corporation (Sigma, São Paulo, Brazil). The transformed tissues were selected on shoot induction medium (4.3 g L^{-1} MS salts, 3% sucrose, 0.8% agar, 1 mL L^{-1} 1000× MS vitamins stock (1000× MS stock solution: thiamine HCl 1 g L^{-1}, nicotinic acid 0.5 g L^{-1}, pyridoxine HCl 0.5 g L^{-1}, glycine 2 g L^{-1}, and myo-inositol 100 g L^{-1}), 0.5 µM µ-naphthaleneacetic acid, 4 µM 6-benzylaminepurine, 100 mg L^{-1} tioxin, and 3 mg L^{-1} phosphinothricin (PPT). The explants were transferred to new selection medium on a weekly basis and the shoots regenerated within three to four weeks. The transformed shoots rooted in two to three weeks and were maintained on 50 mL of shoot induction medium without phytohormones, but supplemented with 100 mg L^{-1} tioxin and 1 mg L^{-1} PPT, in Majenta GA7 containers (Sigma) under continuous light at 26 ± 2 °C. Then the plants were transferred from the culture media to soil and grown in the greenhouse.

Copy number quantification of the transgenic tobacco events

The copy number was calculated for the five pMCG1005 events (T0) using the formula:$2^{Ct_{reference}\,-Ct_{transgene}}$, which does not require a single-copy reference event (Zhang et al. 2015). After this analysis, a single copy sample was selected as a reference to quantify the relative expression levels using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001).

The PCR reaction solution contained 50 ng of genomic DNA (20 ng mL^{-1}), 1× Fast SYBR Green Master Mix (Applied Biosystems, Thermo-Fisher), 0.6 µM each of the forward and reverse *Axi1* primers (Šubr et al. 2006), and 0.8 mM of the *Bar* primers (*Bar*F: 5'-ACAGCGACCACGCTCTTGA-3 and *Bar*R: 5'-GCTCTACACCCACCTGCTGA-3') in a final volume of 10

mL. A 7500 Fast Real-Time PCR System (Thermo Fisher Scientific) was used for the PCR analysis and the amplification conditions were 95 °C for 20 s, followed by 40 cycles of 95 °C for 3 s, and 60 °C for 30 s.

Segregation test in transgenic tobacco plants

Twenty seeds from each of the five transgenic tobacco T1 pMCG1005 events were germinated in MS medium (Murashige and Skoog 1962) without the selection agent. The genomic DNA was extracted from the seedlings at two and a half months after germination, according to Saghai-Maroof et al. (1984). The DNA was quantified by a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific) and an integrity check was carried out using electrophoresis with 1.0% agarose gel stained with GelRed (Biotium).

The presence of the *Bar* gene was confirmed by PCR using 60 ng of genomic DNA, 0.5 µM of each primer (*35SxÔmega*F: 5'-ACTATCCTTCGCAAGACCCTTC-3' and *Bar*R: 5'-CTTCAGCAGGTGGGTGTAGAG-3'), 1× reaction buffer, 0.6 mM of dNTP, 1U Invitrogen Taq DNA polymerase platinum, and 5% DMSO in a final volume of 20 mL. The cycling conditions were 94 °C for 2 min, followed by 35 cycles of: 94 °C for 20 s, 60 °C for 20 s, and 72 °C for 1 min, with a final extension of 72 °C for 5 min.

A χ^2 test was performed to compare the expected proportions of the transgene copy numbers.

RESULTS AND DISCUSSION

Tobacco is a model plant that is often used in plant-microorganism interaction and gene function studies (Clarkson et al. 2017, Chen et al. 2018, Bai et al. 2019). Real-time PCR was used to estimate the copy number in transgenic tobacco plants using *Axi1* as the endogenous gene and *Bar* as the target gene. Quantitative PCR is based on a comparison between the target gene and the single copy gene in each sample. Therefore, having a characterized single-copy gene is essential to ensure experimental accuracy. In this study, the $2^{-\Delta Ct}$ method was used to select a reference sample that could be utilized in large-scale experiments because it does not require a reference sample. After the first round of analysis, a single copy event was selected as the reference sample for the $2^{-\Delta\Delta Ct}$ method, which was then used to estimate the copy number of the *Bar* gene in the transgenic tobacco genome. Finally, a Mendelian segregation analysis was carried out to check the real-time PCR results and to confirm that there was only one *Bar* gene insertion.

The $2^{-\Delta\Delta Ct}$ method requires an amplification efficiency of between 90 and 100%, and an equivalent efficiency between the endogenous control and the target gene. Therefore, the amplification efficiency was tested and the results showed a 93% efficiency rate. This meant that the *Bar* oligonucleotides could be used in the real-time PCR. Furthermore, the dissociation or melting curve showed the presence of a single product (Figure 1).

The standard curve was created using plasmids with increasing concentrations of the *Axi1* gene fragment (727 bp) that ranged from 100 to 100 000 copies in the tobacco genomic DNA. The tobacco genomic DNA was quantified as
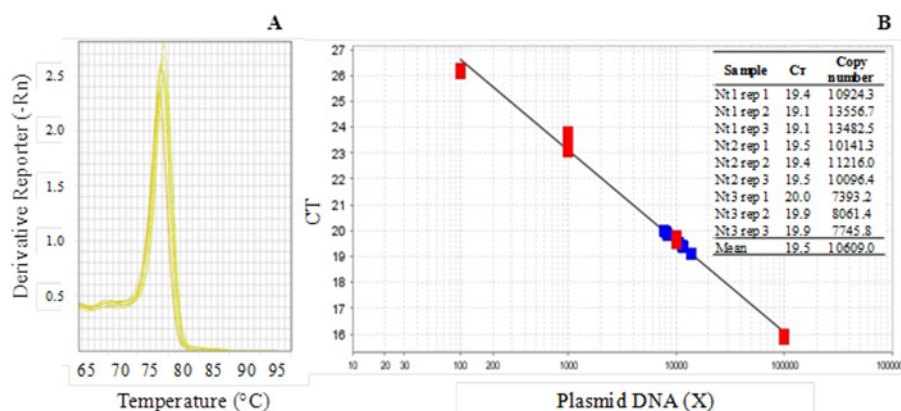

Figure 1. Dissociation curve when *Axi1* primers **(A)** were used. The standard curve for the *Axi1* gene cloned in a plasmid with 100 000, 10 000, 1000, and 100 copies is in red and the *Axi1* gene in the *N. tobacco* genome is in blue. There were three biological and three technical replicates (Nt 1–3, rep 1–3) **(B)**. R²: 0.992 and Efficiency (%): 93.055.

having 7000 and 12 650 copies after interpolating the results from the 10 000 copies point of the standard curve (Figure 1B). This result indicated that there was only one copy of the *Axi1* gene in the tobacco genome and that it was inherited from only one parent.

We generated seven transgenic events (T0), but only five events produced enough seeds to perform the experiments. Initially, the $2^{-\Delta Ct}$ method was performed, using *Axi1* as endogenous the gene and *Bar* as the target gene in the T0 generation of five tobacco transgenic events, to identify a single copy sample because we did not have a characterized single copy sample for tobacco. After the first round of analysis, a reference sample (single copy) was selected and used to calculate the number of copies in the five events. The $2^{-\Delta\Delta Ct}$ method was used to estimate the copy number of the *Bar* gene in the transgenic tobacco genome and then a segregation test was performed (Figure 2). The five analyzed events were considered to be a single copy (Table 1). The values 0.52 and 0.5 were estimated as copy numbers equal to 1 because the transgene was inserted into the T0 generation by hemizygosis (Tt). Although all DNA samples were diluted to 20 ng µL^{-1} and the concentration verified by the Nanodrop method and in agarose gel, which produced similar results to the $2^{-\Delta\Delta Ct}$ method, there was some variation in the pipetting and DNA quality that resulted in different $2^{-\Delta\Delta Ct}$ results and consequently a difference in the copy numbers, which ranged from 0.5 and 1.0. As expected, the wild type tobacco did not show amplification of the *Bar* gene (Table 1). Transgenic single copy events are desirable because they follow a Mendelian segregation pattern, which means that the possibility of gene silencing is lower (Collier et al. 2017, Tark-Dame et al. 2018). Our results corroborated those of Šubr et al. (2006) who also used the *Axi1* gene as an endogenous control to determine the copy number for the Potato virus A (PVA) P3 gene in transgenic tobacco.
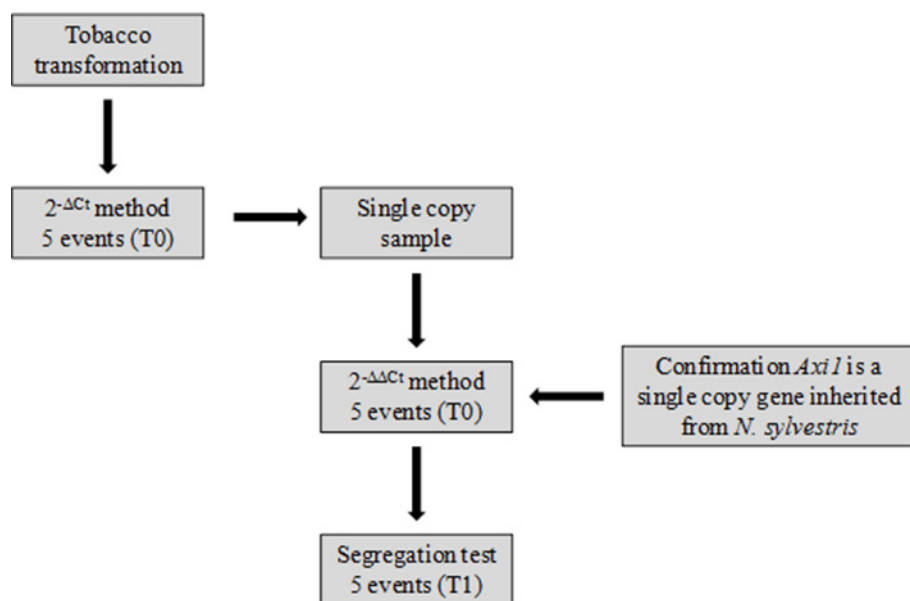

Figure 2. Experimental workflow.

Table 1. Copy number, segregation, and χ^2 results (3481; p < 0.05) for tobacco transgenic event T1

Samples	Copy number ($2^{-\Delta\Delta Ct}$)	Segregation	χ^2
pMCG1005 Ev. 3	0.52	3:1	1.42
pMCG1005 Ev. 4	0.98	3:1	1.06
pMCG1005 Ev. 6	0.5	3:1	0.26
pMCG1005 Ev. 7	1.12	3:1	2.40
pMCG1005 Ev. 10	0.45	3:1	0.26
Wild type tobacco	0	-	-

Once it had been confirmed that there was only one copy of the *Axi1* gene in the *N. tabacum* genome, a PCR was performed with *N. sylvestris* as the maternal parent (S-genome) and *N. tomentosiformis* as the paternal parent (T-genome). The PCR only amplified the *Axi1* gene fragment (975 bp) in the maternal parent (Figure 3).

Nicotiana (Solanaceae) contains approximately 75 species, 48% of which are allotetraploids (36 species) that are classified into 12 sections. These allotetraploids are mainly distributed in the Americas, except for one section that is found outside the Americas (Knnap et al. 2004). *Nicotiana tabacum* (2n = 48, SSTT) was originally derived from interspecific hybridization between *N. sylvestris* (2n = 24, SS) and *N. tomentosiformis* (2n = 24, TT) followed by chromosome doubling. These events occurred less than 1 Ma ago at an early stage in the diploidization process (Clarkson et al. 2010). Concerted evolution of the rDNA has been documented in *N. tabacum*. The process led to rDNA loci being overwritten within a few generations by one dominant progenitor copy, which was usually maternal, and it is possible that the same could have happened to the *Axi1* loci.

N. tabacum genome is highly similar to its parent's genome. However, there has been a reduction in its genome, with greater losses of repetitive DNA sequences from the T-genome than from the S-genome, showing a preferential loss of repetitive sequences from the paternal parent at the genome level (Renny-Byfield et al. 2011). The loss of repetitive sequences from T-genome also happened with *Axi1* gene, which is only present in the S-genome. Studies to identify mutations in the N gene, responsible to resistance to mosaic virus and present in the N-chromosome, also showed that this chromosome is closer related to *N. sylvestris*, which is probably derived from the S-genome (Chen et al. 2018), as showed for *Axi1* gene. The genome downsizing is a widespread biological response to polyploidization that is also illustrated in the *Xyloglucan endotransglucosylase/hydrolase* (*XTH*) gene family, which has a higher number of *NtXTHs* genes in *N. tabacum* than in the ancestral donors, but less than the sum of these two donors (Wang et al. 2018).

Antibiotic or herbicide selection is a conventional method that is used to differentiate between hemizygous and homozygous lines if the transgene is a dominant trait and follows the Mendelian first law of segregation (Passricha et al. 2016, Tarafdar et al. 2019). To validate the data obtained for the copy number by real-time PCR, a segregation test was performed on the five independent pMCG1005 events. For event pMCG1005 Ev. 3, 12 seedlings out of 19 were positive for the *Bar* gene and 7 were negative, which was similar to event pMCG1005 Ev. 4, where 13 out of 20 seedlings were positive for the *Bar* gene and 7 were negative, and event pMCG1005 Ev. 6, where 14 out of 20 seedlings were positive for the *Bar* gene and 6 were negative. However, for event pMCG1005 Ev. 7, 18 out of 20 seedlings were positive for the *Bar* gene and 2 were negative and for event pMCG1005 Ev. 10, 16 out of 20 seedlings were positive for the *Bar* gene and 4 were negative (Table 1). The segregation pattern obtained in these events was the expected for monogenic characteristics (3:1) and this was confirmed by the $\chi2$ test at $p < 0.05$ (Table 1). Events that had copy numbers between 0.5 and 1.0 may have had more than one copy integrated into the same locus. In this case, the copies would be inherited together and show a 3:1 distribution in the segregation test at T1. Tizaoui and Kchouk (2012) reported a tobacco transgenic event that showed two tightly linked inserts in the cis part at T1 with a low recombination frequency. By the T2 generation, the transgenic locus had evolved and the inserts were far enough apart to recombine with a high frequency. Transgenes are sexually inherited as a dominant trait with 3:1 Mendelian inheritance when present as single copy in the host genome
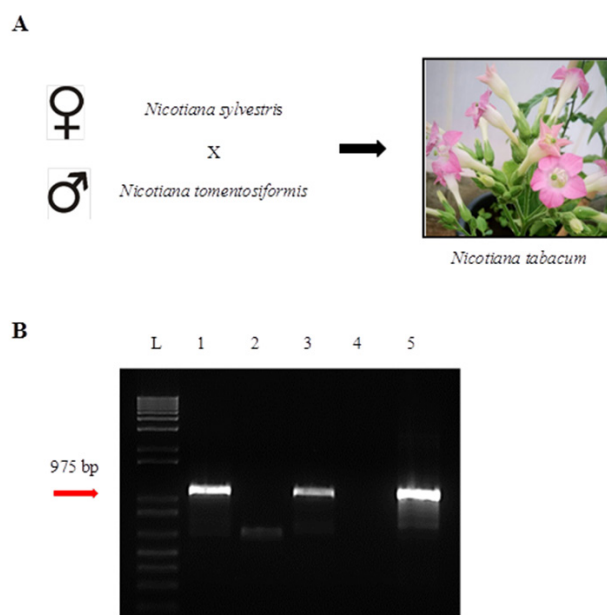

Figure 3. A. Representation of the interspecific hybridization that occurred 400 000–800 000 years ago between *N. sylvestris* and *N. tomentosiformis,* which generated *N. tabacum.* **B.** PCR analysis of the tobacco *Axi1* gene origin. **L**. 1 kb Plus DNA Ladder (Invitrogen), **1**. *N. sylvestris*, **2**. *N. tomentosiformis*, **3**. *N. tabacum*, **4**. Negative control (Water); **5**. Positive control (pGEM®-T *Easy* Vector with the *Axi1* gene). Red arrow indicates the amplified *Axi1* fragment (975 pb).

(Low et al. 2018), which is usually functional (Parrott 2010).

A Southern blot analysis is traditionally used to determine transgene copy number and is still a very useful technique when working with a species that do not have an available, fully annotated genome sequence. While reliable, this classic molecular biology technique is laborious, time-consuming, and requires large amounts of starting plant material, which is difficult in the case of tobacco seedlings, for example, and may involve the use of hazardous radioisotopes (Giancaspro et al. 2017, Nazir et al. 2019). Furthermore, it may fail to detect the exact number of transgene copies when these have been rearranged during integration into the host genome, which may cause changes and/or the loss of relevant restriction sites (Gadaleta et al. 2011, Giancaspro et al. 2017). Real-time PCR analyses are being used on many crops and model species to detect and characterize transgene locus structures. The determination of the transgenic locus number through real-time PCR overcomes the problems linked to phenotypic segregation analysis (i.e., lack of detectable expression even when the transgenes are present) and it is possible to analyze hundreds of samples per day. These two advantages mean that real-time PCR is an efficient method for estimating gene copy numbers (Gadaleta et al. 2011, Thu et al. 2016, Giancaspro et al. 2017). Real-time PCR is a fast and sensitive method, but its precision depends on a reliable one-copy reference gene, which was confirmed by the results from this study.

To date, there have been no genes characterized and validated in tobacco that could be used as a single-copy reference. However, this study has shown that *Axi1* is definitely a single-copy gene in *N. tabacum*, and that it came from its maternal donor *N. sylvestris*. The results also showed that it can be used as a reference gene for copy number determination by real-time PCR and the $2^{-\Delta\Delta Ct}$ method, and consequently for transgenic characterization.

ACKNOWLEDGEMENTS

REFERENCES

Bai G, Yang DA , Cao P, Yao H, Zhang Y , Chen X, Xiao B, Li F, Wang ZY, Yang J and Xie H (2019) Genome-wide identification, gene structure and expression analysis of the MADS-Box gene family indicate their function in the development of tobacco (*Nicotiana tabacum L.*). **International Journal of Molecular Sciences 20**: 1-20.

Burnett MJB and Burnett AC (2019) Therapeutic recombinant protein production in plants: Challenges and opportunities. **Plants, People, Planet 2**: 121-132.

Chen S, Ren F, Zhang L, Liu Y, Chen X, Li Y, Zhang L, Zhu B, Zeng P, Li Z, Larkin RM and Kuang H (2018) Unstable allotetraploid tobacco genome due to frequent homologous recombination, segmental deletion, and chromosome loss. **Molecular Plant 11**: 914-927.

Clarkson JJ, Kelly LJ, Leitch AR, Knapp S and Chase MW (2010) Nuclear glutamine synthetase evolution in Nicotiana: Phylogenetics and the origins of allotetraploid and homoploid (diploid) hybrids. **Molecular Phylogenetics and Evolution 55**: 99-112.

Clarkson JJ, Dodsworth S and Chase MW (2017) Time-calibrated phylogenetic trees establish a lag between polyploidisation and diversification in *Nicotiana* (Solanaceae*).* **Plant Systematics and Evolution 303**: 1001-1012.

Collier R, Dasgupta K, Xing Y-P, Hernandez BT, Shao M, Rohozinski D, Kovak E, Lin J, Oliveira MLP, Stover E, McCue KF, Harmon FG, Blechl A, Thomson JG and Thilmony R (2017) Accurate measurement of transgene copy number in crop plants using droplet digital PCR. **The Plant Journal 90**: 1014-1025.

Edwards KD, Fernandez-Pozo N, Drake-Stowe K, Humphry M, Evans AD, Bombarely A, Allen F, Hurst R, White B, Kernodle SP, Bromley JR, Sanchez-Tamburrino JP, Lewis RS and Mueller LA (2017) A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. **BMC Genomics 18**: 1-14.

Fraley RT, Rogers SG, Horsch RB, Sanders PR, Flick JS, Adams SP, Bittner ML, Brand LA, Fink CL, Fry JS, Galluppi GR, Goldberg SB, Hoffmann NL and Woo SC (1983) Expression of bacterial genes in plant cells. **Proceedings of the National Academy of Sciences of the United States of America 80**: 4803-4807.

Gadaleta A, Giancaspro A, Cardone MF and Blanco A (2011) Real time PCR for detection of precise transgene copy number in durum wheat. **Cellular & Molecular Biology Letters 16**: 652.

Giancaspro A, Gadaleta A and Blanco A (2017) Real-Time PCR for the detection of precise transgene copy number in wheat. In Bhalla PL and Singh MB (eds) **Wheat biotechnology: methods and protocols, methods in molecular biology**. Human Press, New York, p. 251-257.

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG and Fraley RT (1985) A simple and general method for transferring genes into plants. **Science 227**: 1229-1231.

Ingham DJ, Beer S, Money S and Hansen G (2001) Quantitative real time PCR assay for determining transgene copy number in transformed plants. **Biotechniques 31**: 132-141.

Jube S and Borthakur D (2007) Expression of bacterial genes in transgenic tobacco: methods, applications and future prospects. **Electronic Journal of Biotechnology 10**: 452-467.

Knapp S, Chase MW and Clarkson JJ (2004) Nomenclatural changes and a new sectional classification in Nicotiana (Solanaceae). **Taxon 53**: 73-82.

Kochetov AV and Shumny VK (2017) Transgenic plants as genetic models for studying functions of plant genes. **Russian Journal of Genetics: Applied Research 4**: 421-427.

Livak K and Schmittgen TD (2001) Analysis of relative gene expression data using real time quantitative pCR and the $2^{-\Delta\Delta Ct}$. **Methods 25**: 402-408.

Low LY, Yang SK, Andrew Kok DX, Ong-Abdullah J, Tan NP and Lai KS (2018) Transgenic plants: Gene constructs, vector and transformation method. In Çelik Ö (ed) **New visions in plant science.** Intech Open, London, p. 41-61.

McGinnis K, Chandler V, Cone K, Kaeppler H, Kaeppler S, Kerschen A, Pikaard C, Richards E, Sidorenko L, Smith T, Springer N and Wulan T (2005) Transgene-induced RNA interference as a tool for plant functional genomics. **Methods in Enzymology 392**: 1-24.

Masek T, Smykal P, Janotova II, Honys D, Capkova VV and Pechan PM (2000) Isolation of a *Brassica napus* L. cDNA encoding a putative high-mobility-group HMG I/Y protein. **Plant Science 159**: 197-204.

Moon KB, Park JS, Park YI, Song YJ, Lee HJ, Cho HS, Jeon JH and Kim HS (2020) Development of systems for the production of plant-derived biopharmaceuticals. **Plants 9**: 1-21.

Murashige T and Skoog FA (1962) Revised medium for rapid growth and bioassays with tobacco tissue culture. **Physiologia Plantarum 15**: 473-497.

Nazir S, Zaffar Iqbal M and Sajid-ur-Rahman (2019) Molecular identification of genetically modified crops for biosafety and legitimacy of transgenes. In Chen YC (ed) **Gene editing - technologies and applications.** Intech Open, London, p. 1-18.

Nováková S, Mazúrová Ľ, Čeřovská N and Šubr ZW (2005) Transgenic tobacco plants carrying the non-structural P3 gene of potato virus A. **Biologia Plantarum 49**: 593-598.

Parrott W (2010) Genetically modified myths and realities. **New Biotechnology 27**: 545-551.

Passricha N, Saifi S, Khatodia S and Tuteja N (2016) Assessing zygosity in progeny of transgenic plants: current methods and perspectives. **Journal of Biological Methods 3**: e46.

Rasmussen R (2001) Quantification on the Light Cycler. In Dietmaier W, Wittwer C and Sivasubramanian N (eds) **Rapid cycle real-time PCR:methods and applications.** Springer, Berlin, p. 21-34.

Renny-Byfield S, Chester M, Kovarik A, Le Comber SC, Grandbastien MA, Deloger M, Nichols RA, Macas J, Novak P and Chase MW (2011) Next generation sequencing reveals genome downsizing in allotetraploid *Nicotiana tabacum*, predominantly through the elimination of paternally derived repetitive DNAs. **Molecular Biology and Evolution 28**: 2843-2854.

Saghai-Maroof MA, Soliman KA, Jorgensen RA and Allard RW (1984) Ribossomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. **Proceedings of the National Academy of Sciences of the United States of America 81**: 8014-8018.

Šubr Z, Nováková S and Drahovská H (2006) Detection of transgene copy number by analysis of the T_1 generation of tobacco plants with introduced *P3* gene of potato virus A. **Acta virologica 50**: 135-138.

Tarafdar A, Vishwakarma H, Gothandapani S, Bhati M, Biswas K, Prakash A, Chaturvedi U, Solanke AU and Padaria JC (2019) A quick, easy and cost-effective *in planta* method to develop direct transformants in wheat. **Biotech 9**: 1-11.

Tark-Dame M, Weber B, de Sain M, Anggoro DT, Bader R, Walmsley A, Oka R and Stam M (2018) Generating transgenic plants with single-copy insertions using BIBAC-GW binary vector. **Journal of Visualized Experiments 133**: 1-13.

Tizaoui K and Kchouk ME (2012) Genetic approaches for studying transgene inheritance and genetic recombination in three successive generations of transformed tobacco. **Genetics and Molecular Biology 35**: 640-649.

Thu LTM, Thuy VTX, Duc LH, Son LV, Ha CH and Mau CH (2016) RNAi-mediated resistance to SMV and BYMV in transgenic tobacco. **Crop Breeding and Applied Biotechnology 16**: 213-218.

Walden R, Hayashi H, Lubenow H, Czaja I and Schell J (1994) Auxin inducibility and developmental expression of *Axi1*: a gene directing auxin independent growth in tobacco protoplasts. **The EMBO Journal 13**: 4729-4736.

Wang M, Xu Z, Ding A and Kong Y (2018) Genome-wide identification and expression profiling analysis of the Xyloglucan Endotransglucosylase/Hydrolase gene family in tobacco (*Nicotiana tabacum L.*). **Genes 9**: 1-22.

Weng H, Yang L, Liu Z, Ding J, Pan A and Zhang D (2005) A novel reference gene, high-mobility-group protein I/Y, can be used in qualitative and real time quantitative PCR detection of transgenic rapeseed cultivars. **Journal of AOAC International 88**: 577-584.

Zhang Y, Liu Y, Zhang J, Wang G, Wang J and Liu J (2015) Assessment of transgene copy number and zygosity of transgenic maize overexpressing *Cry1Ie* gene with SYBR® Green qRT-PCR. **In Vitro Cellular & Developmental Biology – Plant 51**: 125-134.