


Article

The Structure of T-DNA Insertions in Transgenic Tobacco Plants Producing Bovine Interferon-Gamma

Mikhail Burlakovskiy *, Natalia Saveleva, Andrey M. Rumyantsev, Vladislav V. Yemelyanov ,
Marina V. Padkina and Ludmila Lutova

Department of Genetics and Biotechnology, Saint Petersburg State University, Universitetskaya emb. 7/9, 199034 St. Petersburg, Russia; nata.saveljeva@gmail.com (N.S.); rumyantsev-am@mail.ru (A.M.R.); bootika@mail.ru (V.V.Y.); mpadkina@mail.ru (M.V.P.); la.lutova@gmail.com (L.L.)

* Correspondence: m.burlakovskiy@spbu.ru

Abstract: Many of the most modern drugs are of a protein nature and are synthesized by transgenic producer organisms. Bacteria, yeast, or animal cell cultures are commonly used, but plants have a number of advantages—minimal biomass unit cost, animal safety (plants are not attacked by mammalian pathogens), the agricultural scale of production, and the ability to produce complex proteins. A disadvantage of plants may be an unstable level of transgene expression, which depends on the transgene structure and its insertion site. We analyzed the structure of T-DNA inserts in transgenic tobacco plants (*Nicotiana tabacum* L.) belonging to two lines obtained using the same genetic construct but demonstrating different biological activities of the recombinant protein (bovine interferon-gamma). We found that, in one case, T-DNA was integrated into genomic DNA in the region of centromeric repeats, and in the other, into a transcriptionally active region of the genome. It was also found that in one case, the insert has a clustered structure and consists of three copies. Thus, the structure of T-DNA inserts in both lines is not optimal (the optimal structure includes a single copy of the insert located in the active region of the genome). It is desirable to carry out such studies at the early stages of transgenic plants selection.

Keywords: *Nicotiana tabacum*; transgenic plants; agrobacterial transformation; T-DNA insertion; interferon-gamma



Citation: Burlakovskiy, M.; Saveleva, N.; Rumyantsev, A.M.; Yemelyanov, V.V.; Padkina, M.V.; Lutova, L. The Structure of T-DNA Insertions in Transgenic Tobacco Plants Producing Bovine Interferon-Gamma. *Appl. Sci.* **2022**, *12*, 761. <https://doi.org/10.3390/app12020761>

Academic Editor: Jaeyoon Kim

Received: 26 November 2021

Accepted: 10 January 2022

Published: 12 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The trend in the development of modern pharmacology is the transition from the use of small molecules to the use of protein drugs. Almost a third of all pharmaceuticals in development are of a protein nature (for example, antibodies, vaccines, enzymes, hormones, cytokines, and growth factors), and their share continues to grow [1,2]. Unlike small molecules, proteins can only be obtained in the expression system of living organisms. However, their extraction from natural hosts is often expensive, carries a risk of disease transmission, or has ethical problems. The solution is to use transgenic organisms that produce large amounts of foreign (recombinant) proteins. Moreover, the development of genetic engineering made the appearance of protein drugs possible. The recombinant protein market is projected to exceed USD 3.9 billion by 2024 [3], primarily due to therapeutic proteins.

Unfortunately, these figures are due to the price of modern biopharmaceuticals, which are often inaccessible to the vast majority of the world's population. In the US, the average cost of protein drugs is 25 times higher than low molecular weight preparations. Although recombinant insulin has been on the market for five decades, its cost is still high. This is due to the high cost of equipment (fermenters) combined with the need to isolate, purify, store and transport the target substance under sterile and low-temperature conditions [4].

Most of these disadvantages can be eliminated by the production of protein preparations using transgenic plant producers. The key features of plants in comparison with other

systems of protein synthesis (bacteria, yeast, cell cultures, and whole animal organisms) are infectious safety for humans (plants are not infected with human viruses and prions), extremely low cost of cultivation, as well as the possibility of oral administration preparations in the form of edible producer plants. Oral administration eliminates the most difficult and expensive procedures of isolating and purifying the target protein (accounting for up to 80% of the costs in the production of protein preparations) [5,6]. Plant cells surrounded by a cellulose cell wall can provide long-term storage of the recombinant protein and its protection from the acidic environment of the stomach [4].

However, the production of mammalian proteins in plants is associated with a number of difficulties. A feature of transgenic plants is the possible loss of the activity of the target gene, even successfully introduced into the genome and controlled with a strong promoter. The phenomenon of silencing, which means loss of expression of transferred genes in transgenic plants, was discovered in the early 1990s. The study of this phenomenon showed that the frequency of inactivation of transgenes depends on the number of their copies integrated into the plant genome, as well as the peculiarities of the organization of the insertion (the presence of duplications, vector fragments, etc.) and the site of insertion [7]. Loss of expression can occur at the transcriptional or posttranscriptional stage, in most cases with the participation of small interfering RNAs (miRNAs) [8]. This phenomenon is of great importance since, during the industrial production of a recombinant protein, it is necessary to achieve a stable level of its content in the tissues of the producer organism.

Transgenic tobacco plants synthesizing bovine interferon-gamma were previously obtained in the laboratory of genetic and cellular engineering of plants, Department of Genetics and Biotechnology, St. Petersburg State University [9,10]. Tobacco was used as a model to evaluate the possibility of obtaining immunomodulatory plants for use in animal husbandry. Two lines were selected (B6 and 311), which demonstrated stable inheritance and expression of the introduced gene, as well as the presence of interferon-gamma protein in tissues. The antiviral activity of plant-made interferon was confirmed in bovine cell culture [6,11]. In an experiment on mice, an extract of transgenic plants was shown to be effective when administered orally [11]. It was also found that extracts of recombinant interferon-gamma isolated from equal amounts of tissue of 311 and B6 lines had different antiviral activities. Both lines were obtained as a result of transformation with the same transgenic construct, and the search for the cause of these differences became the goal of this work.

2. Materials and Methods

2.1. Plants

We used tobacco plants (*Nicotiana tabacum* L., $2n = 48$) of the Trabzon variety. Inter311 and InterB lines, carrying the *sIFNG* gene under the control of the 35S promoter and the selective kanamycin resistance gene, were obtained by us earlier at the Department of Genetics and Biotechnology, St. Petersburg State University. Each generation of plants was analyzed for the presence of the transgenic insert and its activity. Both lines showed monogenic inheritance [9,11].

Plants were grown in 100 mL cups on Murashige-Skoog (MS) medium [12], in vitro and in vivo in pot culture at 21–23 °C, 16 h photoperiod. Tobacco seeds were sterilized for 3–5 min in a mixture of 30% hydrogen peroxide and 96% alcohol 1:1, dried on sterile filter paper, and planted in Petri dishes on MS medium with a sugar content of 20 g/L for germination for 5–7 days. Subsequently, sterile seedlings were transferred into cups on the same medium, where they were maintained by microcutting. In order to obtain seeds, the plants were planted in pot culture in a greenhouse.

2.2. PCR

DNA was isolated from plant tissues with the DNeasy plant mini kit (Qiagen, Venlo, The Netherlands) according to the manufacturer's instructions.

Polymerase chain reactions were performed using Taq polymerase, buffer, and nucleotides manufactured by Evrogen (Moscow, Russia). For cloning fragments over 1000 bp, we used the highly efficient Phusion polymerase (Thermo Fisher Scientific, Waltham, MA, USA).

2.3. RT-PCR

RNA was isolated from plant tissues using a Purezol reagent (Bio-Rad, Hercules, CA, USA) according to the manufacturer's instructions. RNA was purified from genomic DNA with the enzyme DNase I, which, in turn, was removed using the RapidOut DNA Removal Kit (Thermo Scientific, Waltham, MA, USA). An amount of 1000 ng of RNA was taken for the reverse transcription in different experiments. Reverse transcription of RNA was performed using RevertAid reverse transcriptase (Thermo Scientific, Waltham, MA, USA) and oligo-dT18 primer according to the manufacturer's instructions. The resulting cDNA samples were diluted with sterile deionized water to a final volume of 100 μ L.

Real-time PCR (RT-PCR) was performed using the Eva Green dye kit (Syntol, Moscow, Russia) in a thermocycler CFX96 Real-Time PCR Detection System (Bio-Rad). Threshold cycles (Ct) were calculated using the CFX-Manager software (Bio-Rad). The quantitative assessment of the expression of the analyzed gene was carried out according to the $2^{-\Delta\Delta C_t}$ method. We used primers to the bovine interferon-gamma gene and to the widely used housekeeping *EF-1a* reference gene (Table 1). Data are presented in relative units, calculated relative to the gene expression level in a calibrator sample.

2.4. Proteins Isolation

Plant material for protein isolation was taken from wild-type plants and from transgenic plants 311.2.7.2- (1–6) and B6.13.8- (1–12) of generation T4. If the mass of one plant grown in vitro was insufficient to obtain the required amount of interferon, samples were collected from up to 12 plants, bringing the total mass of green tissues to 10 g. The numbers of individual plants used are indicated in brackets. Proteins were extracted from young leaves. Raw plant material was homogenized in a chilled porcelain mortar with the addition of chilled extraction buffer (100 mM MES, 1 mM EDTA, 1% glycerol, 50 mM sucrose, 10 mM dithiothreitol, 11 mM ascorbic acid, 1 mM PMSF, pH 7.5, 2 mL per 1 g of plant material). Cross-linked polyvinylpyrrolidone was added to each sample (0.1% by weight of plant tissue). The homogenate was filtered and centrifuged at 10,000 rpm at 4 °C for 15 min, and the supernatant was transferred to a new test tube. For purification of recombinant interferon-gamma from total protein extract, the method of salting-out with ammonium sulfate was used. Interferon protein precipitated in a solution of ammonium sulfate \approx 45%; therefore, initially, the concentration adjusted to 40% of saturation, the samples were left for 1 h with a stirrer at 4 °C, centrifuged for 15 min at 10,000 rpm, the precipitate was removed, and ammonium sulfate was added to the supernatant up to 50% of saturation. Then the solution was centrifuged again at the same conditions, and the supernatant was removed. The precipitate containing IFN-c was dissolved in a buffer suitable for the experiment. Dialysis was used to purify the target protein from low molecular weight impurities, primarily ammonium sulfate. In order to measure the IFN activity in bovine cell culture, dialysis was carried out in a phosphate-buffered saline (PBS), pH 7.4 at 4 °C overnight [11]. A 100 μ L aliquot was taken from the obtained samples for Western blot analysis; the remaining volume (about 1 mL) was tested on a bovine trachea cell culture.

Table 1. Primers used in the process.

T-DNA sequencing primers	p35S (cauliflower mosaic virus promoter)	35S_F	CCTCCTCGGATTCCATTGCCAG
		35S_R	GTCCATCTTTGGGACCACTGTCCG
	OCST (octopine synthase gene terminator)	OCS_F	GCGAGACGCCTATGATCGCATGAT
		OCS_R	GAAACCGGCGGTAAGGATCTGAGC
	pNOS (nopaline synthase gene promoter)	NOS_F	CGATGACGCGGGACAAGCCGT
		NOS_R	GACCTTAGGCGACTTTTGAACGCGC
<i>NPT II</i> (neomycin phosphotransferase)	NPT_F	GTGTTCCGGCTGTCAGCGCAGG	
	NPT_R	CGCAAGGAACGCCCGTCGTGG	
RB and LB (right and left borders of T-DNA)	RB_in_1	GTTTACCCGCCAATATATCCTGT	
	LB_in_1	TGGCAGGATATATTGTGGTGTA	
RT-PCR	<i>EF-1a</i> (reference gene)	EF1	CAAGCGGTCATTCAAGTATGC
		EF2	TGTCCAGGACGATCAATCACA
	<i>RTsIFNG</i> (recombinant bovine gamma-interferon)	RTsIFNG-1	AGGAGTATGGACATCATCAAGCA
		RTsIFNG-2	AGTCGTCGACCGGAATTGA
FPNI-PCR «Genome walking»	T-DNA sequence near right border (RB)	35Sm_antiSP	CTGATCATGAGCGGAGAATTAAGGGA
		35Sm_SP3	TCGGGAAACCTGTCTGTGCCA
		35Sm_SP2	GCTCACTGCCCGCTTTCCAG
		35Sm_SP1	AGTGTAAGCCTGGGGTGCCT
	T-DNA DNA sequence near left border (LB)	NOST_antiSP	CCTAGTTTGC GCGCTATATTTTGT TTC
		NOST_SP3	GAAGCAGATCGTTCAAACATTTGGC
		NOST_SP2	TCGTTTTGGTGCTACCCACGTT
		NOST_SP1	CCA ACTGGCAAATCATCCAGCGT
Confirmation of T-DNA insertion sites in the plant genome	B6 line	B6_R1	TAGCCAACATACCTCAATTGAGTTCCT
		B6_L2	AATGGCCGCAGAGTGAAGCA
	311 line	311_L1	AGCACTTCCCAGAGCGCATATTACC
		311_R1	CAAATGAAGTTTTGATGCTAATGTATGGGAAT

2.5. Western Blot

Separation of proteins for Western blot analysis was performed by denaturing polyacrylamide gel electrophoresis according to the Laemmli method [13]. Plant extracts were leveled in protein content; 20 µg of total protein was applied to the lane. The transfer of proteins from the gel to a nitrocellulose membrane was carried out in a semi-dry blot system (Trans-Blot SD semi-dry transfer cell, Bio-rad, Hercules, CA, USA) at a current of 2 mA/cm² for 30 min. We used primary polyclonal goat antibodies to recombinant interferon-gamma (R&D systems, Abingdon, UK) and secondary polyclonal anti-goat rabbit antibodies labeled with horseradish peroxidase (Agrisera, Vännäs, Sweden). The development of the nitrocellulose membrane was carried out using 3,3'-diaminobenzidine. The color intensity and the area of interferon-gamma spots on the membrane were assessed using the ImageJ 1.41 software. The content of interferon was calculated in relative units, which were obtained by multiplying the area of the spot on the blot by the intensity of its color, and then the approximate content in µg was calculated relative to the commercial preparation of interferon-gamma, which was applied at a concentration of 1 µg per lane (positive control).

2.6. Measurement of Antiviral Activity

The determination of the specific antiviral activity of the interferon-gamma preparation was carried out using a 2–3 day old monolayer of a continuous FBT cell line (fetal bovine tracheal fibroblasts). The antiviral activity of the test interferon preparation was measured in comparison with the industry-standard sample (ISS) of human leukocyte interferon alpha-2, the activity of which is 1000 IU/mL (OSO 871211, Microgene, Moscow, Russia) [14–16].

2.7. Genome Walking

Fusion primer and nested integrated PCR method (FPNI-PCR) [17] was used for sequencing the genomic DNA regions flanking the known sequence of the T-DNA insertion.

For FPNI-PCR, a set of primers FP1–9 was used, consisting of a semi-random region that binds to random sites of genomic DNA and a constant region complementary to the primers FSP1 and FSP2 [17]. Primers SP1–3 were designed to span the known T-DNA sequences (Table 1). The method includes three stages of PCR using Taq polymerase.

The amplified fragments were sequenced at the Research Park “Center for Molecular and Cell Technologies” in St. Petersburg State University. The obtained sequences were aligned using the BLAST NCBI tool and the ApE program.

3. Results

3.1. Evaluation of the Biological Activity of Recombinant Interferon-Gamma in Bovine Cell Culture

The measurement of the biological activity of plant-made interferon-gamma was carried out on a culture of bovine tracheal fibroblasts. We studied extracts of non-transgenic plants (wild type), transgenic plants of 311 and B6 lines of the 4th generation, as well as an extract from frozen leaves of the 311 line (Table 2). Extracts were prepared from an equal mass of plant tissue; the presence of the recombinant protein was confirmed by Western blot (Figure 1).

It was found that the extracts of transgenic plants contained bovine interferon-gamma (Figure 1) in a similar amount in both transgenic lines (463 relative units per lane, which corresponds to 0.185 µg of interferon). The plant extract of the 311 line had antiviral activity comparable to that of gamma-interferon produced in transgenic microorganisms (Table 2). The product from the B6 line was unexpectedly less effective. The weak antiviral activity of the wild-type tobacco extract was found. This may be due to the stimulating effect of the tobacco’s own proteins, which are not completely removed by the used purification methods. The extract of transgenic plants subjected to long-term storage under freezing conditions did not differ in the content of the target protein from other lines (Figure 1 and Table 2), but at the same time, it lost its antiviral activity. Thus, the presence of interferon-gamma protein in the sample does not guarantee the preservation of its antiviral activity.

Table 2. The results of measuring the antiviral activity of interferon preparations.

Test Drug	Activity in IU/mL (5 June 2013)	Activity in IU/mL (13 March 2014)	Activity in IU/mL (21 April 2016)	Activity in IU/mL (20 May 2016)
Negative control (extraction buffer)	0	0	0	0
Extract of non-transgenic tobacco plants	40	0	0	0
Extract of transgenic tobacco plants mix 311.2.7.2 (1–6)	320	320	120	160

Table 2. Cont.

Test Drug	Activity in IU/mL (5 June 2013)	Activity in IU/mL (13 March 2014)	Activity in IU/mL (21 April 2016)	Activity in IU/mL (20 May 2016)
Extract of transgenic tobacco plants mix B6.13.8 (1–12)	80	100	0	0
Extract of frozen transgenic tobacco plant 311.2.7 after a year of storage at -80°C	10	-	-	-
Positive control (commercial recombinant interferon-gamma produced in bacteria)	380 IU/ μg	260 IU/ μg	250 IU/ μg	250 IU/ μg
Industry standard sample (ISS) (OSO 871211, Microgene, Russia)	10	10	10	10

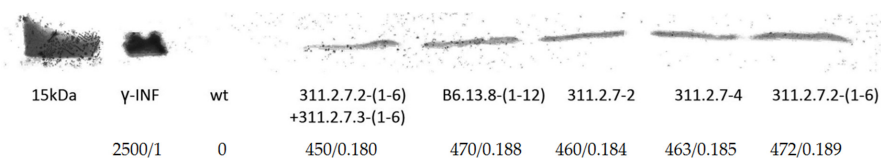


Figure 1. The result of Western blot test of protein samples obtained from T3–4 generations of transgenic plants. γ -INF—commercial bovine interferon-gamma, wt—non-transgenic tobacco, 15 kDa—molecular weight marker. The numbers correspond to extracts from transgenic plants of lines 311 and B6 of different generations. The bottom line presents the results of the semiquantitative determination of the interferon-gamma content on the blot, rel. units/ μg . The sensitivity of our version of Western blotting is probably insufficient for a given protein concentration. This Western blot is best viewed as qualitative analysis for the presence or absence of the target protein in the samples.

3.2. Search for a Possible Reason for the Biological Activity Difference between 311 and B6 Lines

3.2.1. Analysis of sIFNG Gene Expression in Tobacco Plants of Two Lines

The first hypothesis explaining the difference in biological activity between the B6 and 311 lines was associated with a difference in the level of transgene expression. It is important to note that samples for RNA isolation were taken from plants independently of the samples from which the protein was subsequently isolated for the study of biological activity.

In order to test this hypothesis, RT-PCR was carried out with primers for the interferon-gamma gene and the reference gene *EF-1a* (Figure 2). This experiment showed that within each group, there is a significant variation in the level of expression, as a result of which there are no significant differences between the groups (analysis of variance, $p = 0.05$).

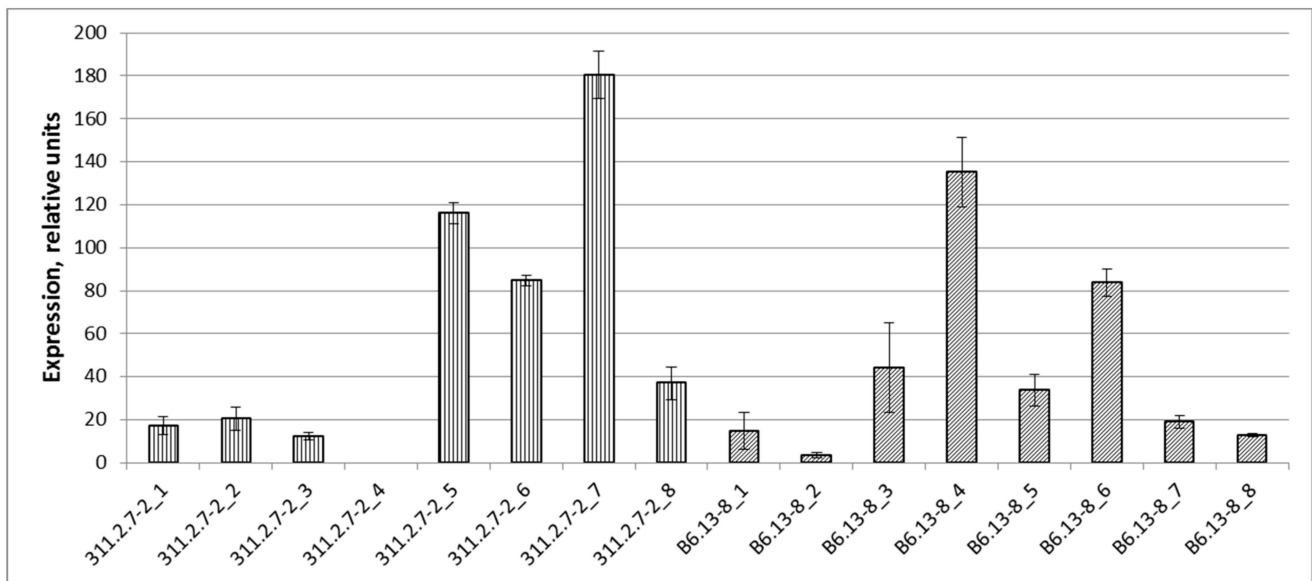


Figure 2. *sIFNG* gene expression in green tissues of individual tobacco plants, lines 311 and b6. Each sample was tested in three technical replicates. Error bars represent standard deviation.

3.2.2. Sequencing T-DNA Insert

The second hypothesis explaining the difference in biological activity between the B6 and 311 lines was due to a mutation in the insertion sequence.

Tobacco plants, the founders of the InterB and Inter311 lines, were transformed with the *Agrobacterium tumefaciens* EHA105 strain carrying the pART27-INT6 plasmid [11,16]. The T-DNA region inserted into the plant genome is limited by short RB and LB (right and left borders) sequences and includes the target *sIFNG* gene under the control of the 35S promoter and the kanamycin resistance gene (encoding neomycin phosphotransferase II) with the nopaline synthase promoter. The nucleotide sequence of plasmid pART27 is not available in the NCBI database, but it is indicated that RB and LB, as well as the 35S promoter and kanamycin resistance gene, were taken from plasmid pGA643 [18], which sequence is known (GenBank: AY804024.1). A pair of primers was selected for each of the functional elements (Table 1) so that the size of the amplified fragment was about 150 base pairs. Thus, using various combinations of primers, the sequence of the transgenic insert was overlapped by fragments of no more than 1000 base pairs (Figure 3).

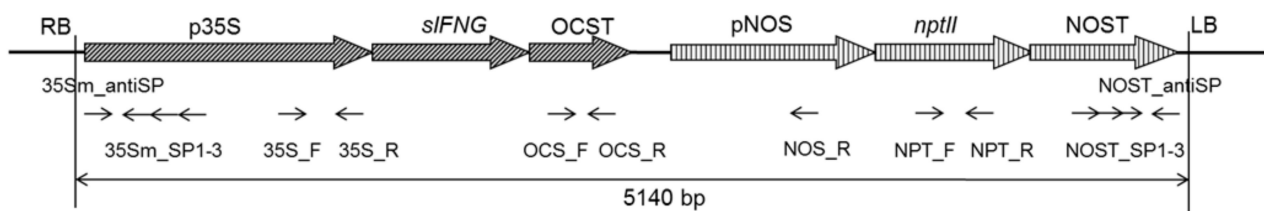


Figure 3. The structure of T-DNA from plasmid pART27-INT6 and primers binding sites. p35S—35S promoter; *sIFNG*—synthetic bovine gamma interferon gene; OCST—octopine synthase terminator; pNOS—nopaline synthase promoter; *nptII*—neomycin phosphotransferase II gene (kanamycin resistance); NOST—nopaline synthase terminator.

Sequencing showed that transgenic inserts in plants 311.2.7.2-2 and B6.13.8-17 (4th generation) were 98–99% identical to each other and the corresponding T-DNA region of plasmids from the databases. The observed mismatches are highly likely to be sequencing artifacts. Mismatches are recognized as artifacts since they are located in the end regions of the Sanger sequences and do not repeat in case of overlapping sequences. For plants

of 311 line, a slight rearrangement (deletion) in the terminal region of the NOS terminator was found.

Thus, no differences were found in the T-DNA sequence between plants of the Inter311 and InterB6 lines, which could explain the difference in the biological activity of their extracts.

3.2.3. Search for T-DNA Insertion Sites

The next hypothesis explains the differences between the InterB and Inter311 lines in the production of the recombinant protein by the transgene integration into regions of the genome with different transcriptional activity.

For sequencing of the tobacco genomic DNA regions flanking the T-DNA insert, the FPNI-PCR “genome walking” method was used. Primers were picked for the T-DNA regions close to the right and left borders of the previously sequenced inserts (Table 1). As a result, fragments up to 968 bp in size were obtained, including sequences of tobacco genomic DNA up to 750 bp.

By using BLAST, it was found that T-DNA in the case of the B6 line was integrated into the centromeric region of the indeterminate chromosome, among repeats and retrotransposons (Figure 4A). The obtained sequences of genomic DNA fragments correspond to regions of the sequence NW_015800096.1 (NCBI Reference Sequence).

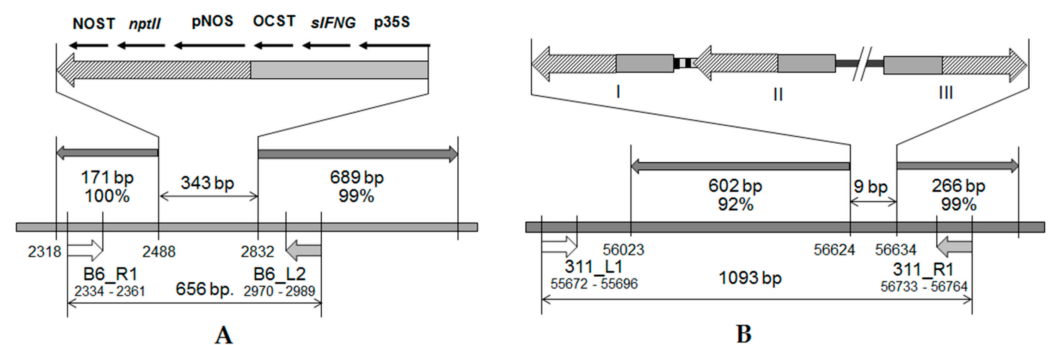


Figure 4. T-DNA insert structures. The figure shows the location of the sequences obtained as a result of genome walking, the length of the overlapping regions, and the percentage of similarity are indicated. The binding sites of the primers used to confirm the genome walking data, and the length of the resulting PCR fragment are also indicated. (A): Insert structure from line B6 (into sequence NW_015800096.1). (B): Insert structure from line 311 (into sequence NW_015801991.1). Numbers I, II, III denote copies of T-DNA.

For 311 line, insertion also occurred in an undefined region of the genome, but in a more active one, next to it the genes “*Nicotiana tabacum* 60S ribosomal protein L39-like” (LOC107799127), “eukaryotic translation initiation factor 3 subunit M-like” (LOC107799143), receptor-like protein kinase FERONIA (LOC104086005), and chloroplastic methylesterase 11 (LOC107799161) are located. Insertion did not destroy any of the open reading frames, i.e., no T-DNA insertional mutagenesis event occurred [19].

Moreover, as a result of genome walking, it was found that 311 line carries a cluster of at least three insertions in a tail-to-tail-head-to-tail orientation (Figure 4B). PCR with specific primers (including a number of PCR reactions using Phusion polymerase to obtain large fragments) confirms that only NOS terminator sequences bind to genomic DNA, and there is a head-to-tail connection in the genome. A difference in the sizes of the insertion sites was found—in the 311 line, T-DNA displaced only nine nucleotides of genomic DNA, while 344 nucleotides of the original sequence disappeared in the B6 line.

In order to confirm the correctness of the determination of the insertion sites, primers were designed for the genomic DNA sequences flanking the insertion sites of both plant lines (Table 1). PCR with primers for line 311 showed that the corresponding fragment (1093 bp, Figure 4B) appeared in wild-type and B6 samples but was absent in sample 311,

in which a large region (T-DNA) was inserted between the binding sites of these primers (Figure 5A). It confirms the correctness of the localization of the insertion site, as well as the fact that no unmodified copies of this region remained in the genome of plants of 311 line. This means that the 311 line is homozygous for the recombinant gene.

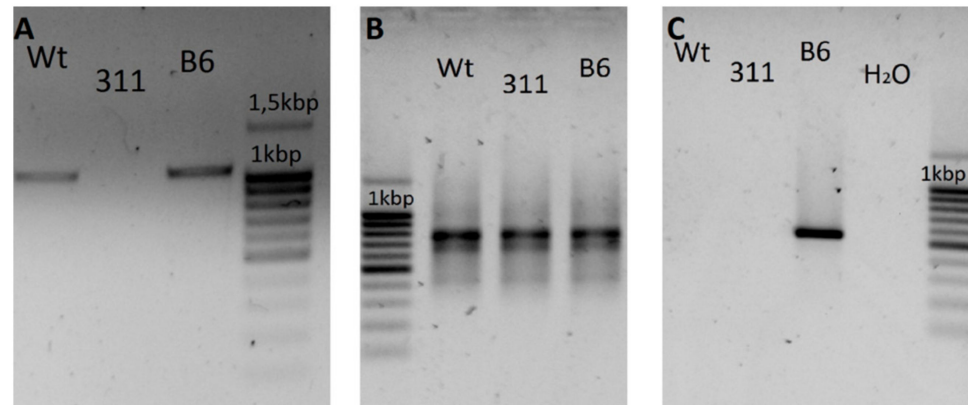


Figure 5. PCR results with primers to the tobacco genomic DNA flanking transgenic inserts. Sample Wt—DNA of non-transgenic tobacco plant, samples 311 and B6—DNA of transgenic plants of lines Inter311 and InterB6. (A): primers 311_L1 + 311_R1. The corresponding fragment (1093 bp) is absent in sample 311, in which the T-DNA was inserted between the binding sites of these primers. (B): primers B6_L2 + B6_R1. High level of non-specific binding was shown for all samples, which confirms the insertion of the T-DNA into the region of repeats. (C): primers B6_R1 (tobacco genomic DNA) + Nost_SP2 (T-DNA region). The amplified fragment appeared only in sample B6.

A similar reaction with primers for the insertion site of the InterB6 line (Figure 4A) showed a high level of non-specific binding for all samples (Figure 5B). This confirms the incorporation of T-DNA into the repeat region, which makes it impossible to accurately localize the insertion site, as in the case of plants of 311 line. PCR with primers to the B6 genomic DNA region (B6_R1) and the T-DNA region (Nost_SP2) gives the best result—an amplified fragment appeared only in a sample with B6 line DNA, which confirms the correctness of the localization (Figure 5C).

3.2.4. The Structure of the Junctions of T-DNA and Genomic DNA

Rearrangements occurred in the terminal part of the nopaline synthase terminator of the 311 line (Figure 6). For B6 plants, the full-length NOS terminator is present, separated from the plant genome by a region of plasmid DNA corresponding to the shortened left border. For the 311 line, any LB fragments are absent at both junction points of T-DNA and plant genomic DNA, and the NOST was shortened by 160–200 bp. The most interesting structure is the site of two inserts joining in the head-to-tail orientation: the 95 bp T-DNA region, including the NOST and LB fragments, was turned 180 degrees (Figure 6). These rearrangements did not affect the functionality of the NOST, judging by the fact that 311 line plants successfully passed the stage of kanamycin selection.

The exact structure of the “tail-to-tail” junction of the inserts remains unknown—this region could not be fully amplified either during FPNI-PCR or standard PCR with 35Sm-SP primers. This indicates a large size of the intermediate section. According to the literature, there are cases of insertion into the genome of a transgenic plant of additional sequences formed by plant and/or plasmid DNA [20].

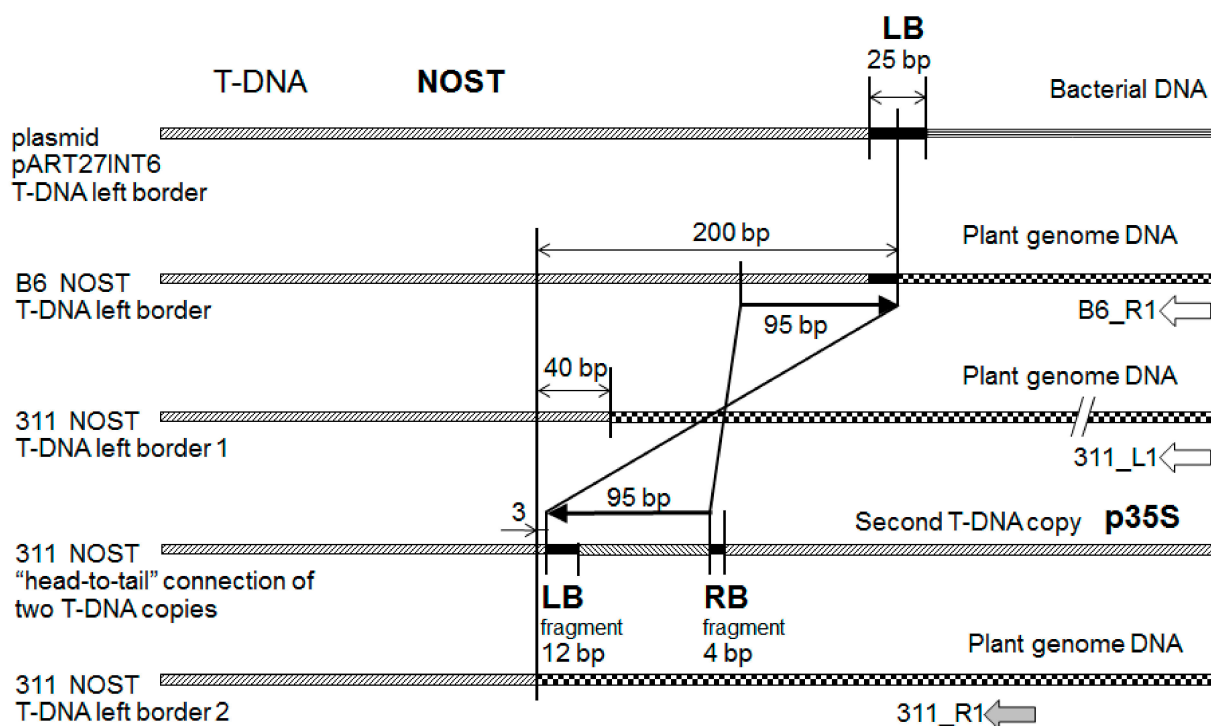


Figure 6. Scheme of T-DNA and genomic DNA junctions, as well as the junction point of inserts from line 311.

4. Discussion

Recombinant interferon-gamma of plant origin showed antiviral activity comparable to the commercial preparation of interferon produced in bacteria (Table 2). Thus, the possibility of synthesis of biologically active bovine interferon-gamma in the tissues of transgenic plants was confirmed. The preparation of plant tissues of 311 line, frozen for storage, lost its antiviral activity (Table 2). This may be due to too strong freezing (-80°C) without the use of cryoprotectants. Most likely, the loss of the activity of the recombinant protein in tissues is associated with a disturbance of its folding since the content of the target protein did not differ from other lines (Figure 1 and Table 2). It was also found that extracts of recombinant interferon-gamma, isolated from tissue samples of an equal weight of transgenic plants 311 and B6 lines, exhibit different biological activity, although Western blot analysis showed approximately the same content of interferon-gamma protein in them (Figure 1). The most likely explanation is the lack of sensitivity of our Western blot version for a given protein concentration. Thus, despite the possibility of quantitative measurement, this Western blot is better considered as qualitative analysis for the presence of the target protein in the samples. Considering that both lines were obtained as a result of transformation with the same transgenic construct, the next important task of the study was to find the cause of these differences.

The first hypothesis explaining the difference in biological activity between the B6 and 311 lines was associated with a difference in the level of transgene expression. It was found that even individuals from the same inbred line after several generations of the selection show a significant difference in the level of transgene expression. This is consistent with the literature data, according to which there is often a significant variation in the expression level in the offspring of one transformant [21]. Such a significant intragroup variation does not allow for a reliable comparison of the two groups. In the final experiment to measure the biological activity of the target protein, this spread was compensated by mixing samples from several (6–12) plants.

The second hypothesis was a mutation in the target gene. This could have happened during the insertion of T-DNA into the plant genome or later. To test this hypothesis, we

decided to sequence the transgenic inserts of both lines. As a result, no differences were found in the T-DNA sequence between plants of 311 and B6 lines, which could explain the difference in the biological activity of their extracts. It should be noted that in the case of line 311 with multiple insertions, the mutation in one of them may be hidden by the suppressive level of the PCR product from the others.

The next hypothesis explains the differences in the productivity of transgenic plants by the integration of transgenes into regions of the genome with different activities. In order to test this hypothesis, localization of the T-DNA insertion sites in both studied lines was required, which was performed by the “genome walking” method (FPNI-PCR). This made it possible to obtain sequences of genomic DNA regions flanking transgenic inserts. Analysis of the obtained sequences showed that, in the case of the 311 line, the insertion actually occurred in the transcriptionally active region of the genome. For the B6 line, an insert was found in the region of centromeric repeats. The results of the genome walking were confirmed by PCR with specific primers (Figure 5). Thus, the data obtained confirmed the hypothesis of T-DNA insertion into genome regions with different transcriptional activity. Integration into an inactive region of the genome can explain the lower biological activity of the extract from B6 line plants. Although early studies showed that T-DNA is predominantly incorporated into the active regions of the genome, more recent studies have confirmed that T-DNA is incorporated equally likely into all regions of the genome [22,23].

The sequences obtained during FPNI-PCR showed that the insert in the 311 line has a complex trimeric structure and consists of three identical copies of T-DNA in a tail-to-tail-head-to-tail orientation. This fact could not be detected by genetic analysis, which demonstrated monogenic inheritance of the transgene. Multi-copy insertion often causes silencing (decrease in transgene expression) [20]. DNA repeats, especially inverted ones, tend to form secondary structures, which usually leads to a weakening of gene expression [24]. However, in this case, the insertion into the transcriptionally active part of the genome was of greater importance, providing a higher recombinant protein activity in the 311 plant line compared to the B6 line. Perhaps the reason is the presence in the tail-to-tail connection of an insert of unknown DNA of significant size, which alienated the inverted copies from each other. Moreover, according to some data, the presence of three copies of transgene in the genome may not affect the level of its expression, and many lines selected for a high level of expression carry three or more copies of the transgene [21].

Complex rearrangements (deletion) were found in the terminal region of the NOS terminator at the insert in the 311 line (Figure 6). According to the literature data, the left border (LB) undergoes more significant rearrangements than the RB when T-DNA is inserted into the host genome [25,26]. In our case, they did not affect the functionality of the target interferon-gamma gene and may be the result of errors in the plant repair system during T-DNA insertion, which led to the cluster structure of the transgene. It should be noted that the missing fragment did not disappear without a trace but remained in the genome at the head-to-tail T-DNA junction point. It is difficult to explain how it turned out to be reversed. Nevertheless, it can be assumed that T-DNA was transported into the cell completely and only then underwent rearrangement in the forehead region. The resulting truncated T-DNA was copied twice by plant DNA synthesis and repair systems, resulting in a cluster of three copies. In general, the results can hardly be called unusual since the insertion of T-DNA into the genome is often accompanied by chromosomal aberrations [20,23,27]. In our case, no residues of vector DNA were found around the inserts, although, according to the literature data, in the case of tobacco plants, 75–80% of transformants contain fragments of vector backbone [25].

The purpose of the experiment was to test the biological activity of the recombinant protein, which is the endpoint in the chain “presence of transgene”-“presence of mRNA”-“presence of protein”-“biological activity of protein”. As practice shows, variations are possible at all stages of this path. Spatial and temporal variations in mRNA, as well as the local availability of resources for protein biosynthesis, strongly influence the relationship between protein levels and their coding transcripts [28]. Western blot detection of the

presence of a protein does not guarantee the preservation of its biological activity due to the peculiarities of folding and protein stability [6]. It is important for the production of recombinant protein for industrial purposes, especially for medical and veterinary purposes.

5. Conclusions

Thus, it was found that transgenic tobacco plants 311 line characterized by increased efficiency in the production of the target protein carries a cluster of three T-DNA inserts in the transcriptionally active region of the genome; the B6 line with a high probability carries one insert in the region of centromeric repeats. Although multiple insertions should have caused silencing and the associated decrease in the level of transgene expression, integration into the active region of the genome had a greater effect. Nevertheless, the structure of transgenic inserts in both lines cannot be called optimal. It is desirable to select plants with a single copy of the transgene in the active region of the genome. Moreover, the assessment of the effectiveness of a transgenic producer organism by the level of gene expression or the presence of a recombinant protein may be incorrect; the endpoint is the measurement of the biological activity of the resulting substance per mass unit of raw material.

Author Contributions: N.S. obtained the transgenic plants, M.B. carried out the experiment with plants and wrote the article, V.V.Y. advised on protein isolation and Western blotting, A.M.R. worked with bovine cell culture, M.V.P. coordinated the experiment with bovine cells, L.L. supervised the project. All authors discussed the results and contributed to the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The article was made with support of the Ministry of Science and Higher Education of the Russian Federation in accordance with agreement № 075-15-2020-922 date 16.11.2020 on providing a grant in the form of subsidies from the Federal budget of Russian Federation. The grant was provided for state support for the creation and development of a World-class Scientific Center “Agrotechnologies for the Future”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The research was performed using the equipment of Research Park “Center for Molecular and Cell Technologies” at Saint-Petersburg State University. We also express our gratitude to Gradoboeva A.E. for her help in working with cell culture.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Walsh, G. Biopharmaceutical Benchmarks 2018. *Nat. Biotechnol.* **2018**, *36*, 1136–1145. [[CrossRef](#)] [[PubMed](#)]
2. Ghag, S.B.; Adki, V.S.; Ganapathi, T.R.; Bapat, V.A. Plant Platforms for Efficient Heterologous Protein Production. *Biotechnol. Bioprocess. Eng.* **2021**, 1–22. [[CrossRef](#)] [[PubMed](#)]
3. Schillberg, S.; Raven, N.; Spiegel, H.; Rasche, S.; Buntru, M. Critical Analysis of the Commercial Potential of Plants for the Production of Recombinant Proteins. *Front. Plant Sci.* **2019**, *10*, 720. [[CrossRef](#)] [[PubMed](#)]
4. Xiao, Y.; Kwon, K.-C.; Hoffman, B.E.; Kamesh, A.; Jones, N.T.; Herzog, R.W.; Daniell, H. Low Cost Delivery of Proteins Bioencapsulated in Plant Cells to Human Non-Immune or Immune Modulatory Cells. *Biomaterials* **2016**, *80*, 68–79. [[CrossRef](#)] [[PubMed](#)]
5. Burnett, M.J.B.; Burnett, A.C. Therapeutic Recombinant Protein Production in Plants: Challenges and Opportunities. *Plants People Planet* **2020**, *2*, 121–132. [[CrossRef](#)]
6. Saveleva, N.V.; Burlakovskiy, M.S.; Yemelyanov, V.V.; Lutova, L.A. Transgenic Plants as Bioreactors to Produce Substances for Medical and Veterinary Uses. *Russ. J. Genet. Appl. Res.* **2016**, *6*, 712–724. [[CrossRef](#)]
7. Gelvin, S.B. Agrobacterium-Mediated Plant Transformation: The Biology behind the “Gene-Jockeying” Tool. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 16–37. [[CrossRef](#)] [[PubMed](#)]
8. Marenkova, T.V.; Deineko, E.V. Transgenic Plants as a Model for Studying Epigenetic Regulation of Gene Expression. *Vavilov J. Genet. Breed.* **2015**, *19*, 545–551. [[CrossRef](#)]

9. Saveljeva, N.V.; Kurdyukov, I.D.; Dudnik, Y.E.; Yemelyanov, V.V.; Padkina, M.V.; Lutova, L.A. Making Plants Producing Bovine Gamma-Interferon for Prophylaxis of Tuberculosis and Leucaemia of Cattle. *Bull. St. Petersburg Univ.* **2009**, *3*, 65–80.
10. Saveleva, N.V.; Lutova, L.A. *Plants—Producers of Recombinant Proteins for Medical Use: Producers of Bovine Interferon γ* ; Lap Lambert Academic Publishing: Saarbrücken, Germany, 2010.
11. Burlakovskiy, M.S.; Saveleva, N.V.; Yemelyanov, V.V.; Padkina, M.V.; Lutova, L.A. Production of Bovine Interferon-Gamma in Transgenic Tobacco Plants. *Plant. Cell Tiss. Organ. Cult.* **2015**, *122*, 685–697. [[CrossRef](#)]
12. Murashige, T.; Skoog, F. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiol. Plant.* **1962**, *15*, 473–497. [[CrossRef](#)]
13. Laemmli, U.K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **1970**, *227*, 680–685. [[CrossRef](#)] [[PubMed](#)]
14. Pontzer, C.H.; Johnson, H.M. Measurement of Interferons. In *Methods in Neurosciences*; Phillips, M.I., Evans, D., Eds.; Neuroimmunology; Academic Press: Cambridge, MA, USA, 1995; Volume 24, pp. 3–9.
15. Padkina, M.V.; Parfenova, L.V.; Gradoboeva, A.E.; Sambuk, E.V. Heterologous Interferons Synthesis in Yeast *Pichia Pastoris*. *Appl. Biochem. Microbiol.* **2010**, *46*, 409–414. [[CrossRef](#)]
16. Lutova, L.A.; Padkina, M.V.; Saveleva, N.V.; Yemelyanov, V.V.; Burlakovskiy, M.S.; Zobnina, A.E.; Tkachenko, A.A. Recombinant Plasmid DNA PART27int6 and Method of Obtaining on Its Basis of Inbred Line of Tobacco Plants Synthesizing Intracellular Gamma-Interferon of Bull. RU2564115C1, 27 September 2015.
17. Wang, Z.; Ye, S.; Li, J.; Zheng, B.; Bao, M.; Ning, G. Fusion Primer and Nested Integrated PCR (FPNI-PCR): A New High-Efficiency Strategy for Rapid Chromosome Walking or Flanking Sequence Cloning. *BMC Biotechnol.* **2011**, *11*, 109. [[CrossRef](#)]
18. Gleave, A.P. A Versatile Binary Vector System with a T-DNA Organisational Structure Conducive to Efficient Integration of Cloned DNA into the Plant Genome. *Plant. Mol. Biol.* **1992**, *20*, 1203–1207. [[CrossRef](#)] [[PubMed](#)]
19. Deineko, E.V.; Zagorskaya, A.A.; Shumny, V.K. T-DNA-Induced Mutations in Transgenic Plants. *Russ. J. Genet.* **2007**, *43*, 1–11. [[CrossRef](#)]
20. Loginova, D.B.; Shumny, V.K.; Deineko, E.V. Features of T-DNA Insert Organization in Transgenic Tobacco Plants Line NU 21. *VOGiS Bull.* **2010**, *14*, 134–140.
21. Kohli, A.; Melendi, P.G.; Abranches, R.; Capell, T.; Stoger, E.; Christou, P. The Quest to Understand the Basis and Mechanisms That Control Expression of Introduced Transgenes in Crop Plants. *Plant. Signal. Behav.* **2006**, *1*, 185–195. [[CrossRef](#)]
22. Kim, S.-I.; Veena; Gelvin, S.B. B. Genome-Wide Analysis of *Agrobacterium* T-DNA Integration Sites in the *Arabidopsis* Genome Generated under Non-Selective Conditions. *Plant J.* **2007**, *51*, 779–791. [[CrossRef](#)]
23. Lacroix, B.; Citovsky, V. Pathways of DNA Transfer to Plants from *Agrobacterium Tumefaciens* and Related Bacterial Species. *Ann. Rev. Phytopathol.* **2019**, *57*, 231–251. [[CrossRef](#)] [[PubMed](#)]
24. Matveeva, T.V.; Sokornova, S.V. Biological Traits of Naturally Transgenic Plants and Their Evolutional Roles. *Russ. J. Plant. Physiol* **2017**, *64*, 635–648. [[CrossRef](#)]
25. Nicolia, A.; Ferradini, N.; Veronesi, F.; Rosellini, D. An Insight into T-DNA Integration Events in *Medicago Sativa*. *Int. J. Mol. Sci.* **2017**, *18*, 1951. [[CrossRef](#)]
26. Singer, K. The Mechanism of T-DNA Integration: Some Major Unresolved Questions. *Curr. Top. Microbiol. Immunol.* **2018**, *418*, 287–317. [[CrossRef](#)] [[PubMed](#)]
27. Kleinboelting, N.; Huep, G.; Appelhagen, I.; Viehoveer, P.; Li, Y.; Weisshaar, B. The Structural Features of Thousands of T-DNA Insertion Sites Are Consistent with a Double-Strand Break Repair-Based Insertion Mechanism. *Mol. Plant* **2015**, *8*, 1651–1664. [[CrossRef](#)] [[PubMed](#)]
28. Liu, Y.; Beyer, A.; Aebersold, R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **2016**, *165*, 535–550. [[CrossRef](#)]