Nucleic Acids Research, Vol. 19, No. 20 5561–5567

Compositional bimodality of the nuclear genome of tobacco

Giorgio Matassi, Roberta Melis, Gabriel Macaya+ and Giorgio Bernardi*
Laboratoire de Génétique Moléculaire, Institut Jacques Monod, 2 Place Jussieu, 75005 Paris, France

Received July 24, 1991; Revised and Accepted September 13, 1991

ABSTRACT

We have studied the compositional distribution of six genes (or small multigene families) and of one family of transposable elements, Tnt1, in DNA fractions from tobacco (*Nicotiana tabacum*) separated according to base composition. We have shown that gene distribution is bimodal and that such bimodality is due to the different base composition of the two parental genomes of tobacco (*N.sylvestris* and *N.tomentosiformis*) and to the different parental origin of the genes tested. These results indicate a physical separation and an absence of extensive recombination of the parental genomes, which have been together in the tobacco nucleus for a small span of their evolutionary life, and a conservation of their compositional patterns, including gene localization.

INTRODUCTION

The compositional approach developed in our laboratory in order to study the organization and evolution of the nuclear genome of eukaryotes has been mainly applied so far to vertebrates (1, 2) and angiosperms (3–5). In both cases, it has been shown that such genomes consist of isochores, long (over 300 kb, on the average, in mammalian genomes), compositionally homogeneous DNA segments, that belong to a number of families characterized by different GC levels; the GC range of isochore families may be different in different classes of animals and plants. The compositional properties of large (over 100 Kb) DNA fragments reflect those of the isochores (from which they derive by the unavoidable enzymatic and mechanical degradation accompanying the preparation of DNA), and define a compositional pattern, or a compositional phenotype of the genome (6). Another compositional pattern is that defined by the compositional properties of genes (exons, introns and individual codon positions; 7).

The compositional patterns, as seen at the DNA fragment (or isochore) level and at the gene level, respectively, are correlated by linear relationships (1, 4, 5, 8). These were demonstrated by localizing genes in DNA fragments fractionated according to their base composition, and by showing that the GC levels of genes (exons, introns and individual codon positions) were linearly correlated with those of the large DNA fragments harboring them. Such gene localization studies not only established the compositional correlations just mentioned, but also provided information on the distribution of genes in the genomes investigated. This revealed that the distribution of genes in the genome of warm-blooded vertebrates is strongly non-uniform, gene concentration being by far the highest in the GC-richest DNA fractions (1, 9). Similar localization studies were performed on plant genomes and led to information on gene distribution (5).

During the course of gene localization experiments on the isochores of tobacco (*Nicotiana tabacum*), we discovered that the distribution of genes in compositional DNA fractions is bimodal. Here, we present these results, which were obtained by using probes for six genes (or small multigene families), glutamine synthase, lignin-forming peroxidase, endochitinase, ribulose bisphosphate carboxylase small subunit (rubisco), nitrate reductase, β-1,3-glucanase, and for one family of tobacco transposable elements, Tnt1. We demonstrate that the bimodal distribution mentioned above is due to a different base composition of the genomes of the two parental species from which tobacco was derived, which are closely related to the present-day species *N.sylvestris* and *N.tomentosiformis*, and to the different parental origin of the genes tested. This provides additional evidence in favor of a physical separation of the two parental genomes in tobacco. These results indicate a conservation of the compositional patterns of the isochores making up the two parental genomes, as well as a conservation of the gene localization in those isochores. Moreover, they suggest the absence of any extensive recombination between two genomes which have been in the same nucleus for a time which has been estimated to be less than 6 million years, whereas the two parental genomes have an estimated divergence time of about 75 million years (10).

MATERIALS AND METHODS

Preparation and fractionation of total DNA

Total DNA was extracted from leaves of wild-type tobacco (*Nicotiana tabacum* cv. Xanthi XHFD8) as described (11), with minor modifications.

Fractionation of DNA (which was at least 50 kb in size) by preparative centrifugation in Cs_2SO_4 density gradient in the

* To whom correspondence should be addressed

+ Permanent address: Universidad de Costa Rica, Ciudad Universitaria Rodrigo Facio, San José, Costa Rica

presence of BAMD, 3,6 bis(acetato-mercuri-methyl) dioxane, and analytical ultracentrifugation of the corresponding DNA fractions in CsCl density gradient were performed as previously described (12; see also ref. 3).

Restriction endonuclease digestion, electrophoresis and hybridization

Unfractionated DNA (6 μg) and DNA fractions (2.5 μg each) were digested with EcoRI (Boehringer, Mannheim, FRG), subjected to electrophoresis on 0.6% agarose gels in TAE buffer (89 mM Tris, 2.5 mM EDTA, 89 mM boric acid, pH 8.0), and blotted onto Hybond N^+ (Amersham, U.K.) nylon membranes by alkaline transfer (unless otherwise indicated). Hybridization with appropriate labelled probes was performed as described in the legends of figures. Autoradiography was performed at −80°C with an intensifying screen.

RESULTS

Compositional fractionation of tobacco DNA

Table 1 presents the buoyant densities in CsCl, the GC levels and the relative amounts of tobacco DNA fractions, as obtained after preparative ultracentrifugation in Cs_2SO_4 in the presence of a DNA ligand, BAMD (see Materials and Methods). Unfractionated tobacco DNA showed a buoyant density of 1.6963 g/cm^3. The results obtained on DNA fractions indicated that buoyant densities ranged from 1.6949 to 1.7016 g/cm^3. The vast majority, 80−90%, of DNA covered, however, an extremely narrow range of about 2 mg/cm^3. This would be equivalent to about 2% GC in the absence of DNA methylation. The buoyant density range corresponded, however, to a much larger spread in base composition, 3.7% GC, owing to the fact that DNA methylation (which causes a decrease in buoyant density; 13−15) increases in compositional fractions as their GC levels increase (Matassi G., Melis R., Kuo K.C., Gehrke C.W. and Bernardi G., paper in preparation). The overall effect is a narrowing down of the buoyant density profile.

Localization of single-copy and small multigene family genes in tobacco DNA fractions

The compositional fractions of tobacco DNA, described in the preceding section, were digested with EcoRI and hybridized with six probes for single-copy genes and small multigene family genes. The results obtained are presented below.

(i) In the case of the nuclear genes encoding chloroplastic glutamine synthase (Fig. 1), two major hybridization bands (7.1 and 11.2 kb), corresponded to the two genes present in tobacco (B. Hirel, personal communication). Experiments on fractionated DNA showed that the 11.2 kb band was localized in fractions 3 and 4, which practically exhibited the same buoyant densities (see Table 1), whereas the 7.1 kb band showed a hybridization peak in fraction 8 (1.6967 g/cm^3; this fraction was slightly overloaded on the gel). Two weaker bands (arrows), showed the same distributions as the 11.2 kb and 7.1 kb bands, respectively.

(ii) The hybridization pattern of the lignin-forming peroxidase gene probe (16) was characterized by four bands, three major ones corresponding to fragments 9.0, 7.5, 6.5 kb in size, and a very weak band corresponding to a fragment of 4.4 kb (Fig. 2). The latter band, although not visible in Fig. 2, was clearly detected both on the original autoradiography and in other experiments. When hybridization was performed on DNA fractions, the 9.0 kb band was centered on fraction 4, and the other bands on fraction 8.

(iii) The hybridization results obtained with the tobacco endochitinase probe (17) showed two sets of bands, which exhibited different distributions in DNA fractions (Fig. 3). One set comprised a 5.7 kb band and a band lower than 0.5 kb in

Table 1. Buoyant densities, GC levels and relative amounts of tobacco DNA fractions from a preparative Cs_2SO_4 /BAMD density gradient centrifugation (a)

Fraction number	ϱ (g/cm^3)	GC (%)	DNA (%)
Total	1.696_3		100
1 (Pellet)	1.694_9	37.3	5.3
2	1.695_6	38.6	2.4
3	1.695_7	38.8	8.0
4	1.695_6	38.6	8.0
5	1.695_7	38.8	10.4
6	1.696_1	39.5	12.1
7	1.696_5	40.3	10.7
8	1.696_7	40.7	13.7
9	1.696_9	41.0	10.6
10	1.697_1	41.4	7.8
11	1.697_8	42.8	6.6
12	1.698_2	43.5	1.8
13	1.699_7	46.3	1.4
14	1.701_6	49.9	1.2

(a) ϱ is the modal buoyant density of the fractions. GC values are from (Matassi G., Melis R., Kuo K.C., Gehrke C.W. and Bernardi G., paper in preparation).

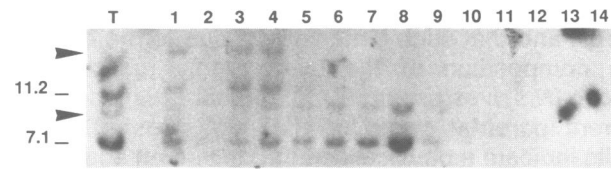

Figure 1. Localization of the chloroplastic glutamine synthase gene. The probe was a 1.5 kb EcoRI fragment of a tobacco cDNA clone, pGS217 (B. Hirel, personal communication). Hybridization was performed in 5×SSPE (20×SSPE is 3M NaCl, 0.2M NaH_2PO_4, 0.02M EDTA, pH=7.7),1% SDS (sodium-dodecyl-sulphate), 0.5% dry milk powder (Boehringer Mannheim, blocking reagent) and 0.1 mg/ml salmon sperm DNA with radiolabeled probe (1×10^6 cpm/ml) at 65°C. Filters were washed for 15 min. in 2×SSC (20×SSC is 3M NaCl, 0.3M trisodium citrate), 0.1% SDS; 15 min. in 0.5×SSC, 0.1% SDS; 15 min. in 0.1×SSC, 0.1% SDS at room temperature, and 10 min. in 0.1×SSC, 1% SDS at 50°C. Faint bands (arrows) were also detected for the two genes. T indicates unfractionated DNA. Molecular weights of the hybridizing bands are given in kb (kilobase pairs).

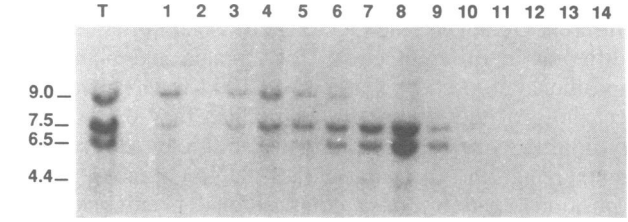

Figure 2. Localization of the lignin-forming peroxidase gene. The probe was a 1.2 kb EcoRI fragment of the tobacco peroxidase cDNA clone, pOD7 (16). Hybridization with radiolabeled probe (1.4×10^6 cpm/ml) was performed as in the experiment of Fig. 2. Filters were washed for 15 min. in 2×SSC, 0.1% SDS at room temperature, and 30 min. in 1×SSC, 1% SDS at 50°C.

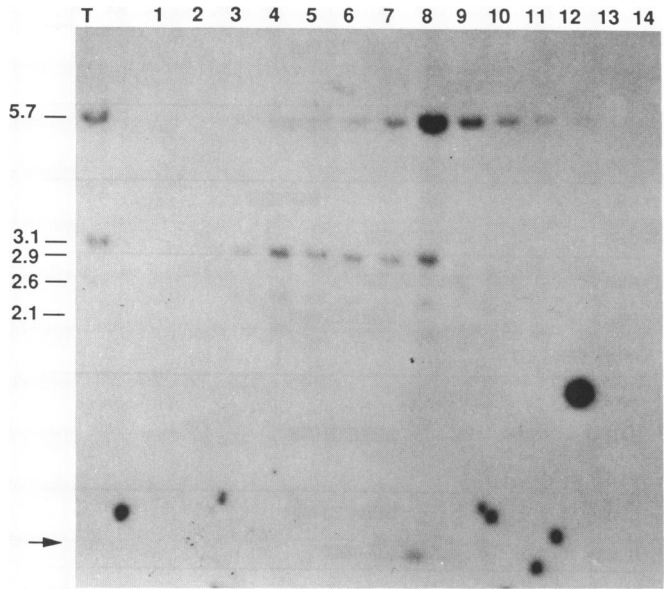

Figure 3. Localization of the endochitinase gene. The probe was the PstI insert of the tobacco endochitinase cDNA clone pCHN50 (17). Radiolabeled probe was added at a concentration of 1.5×10^6 cpm/ml. Filters were washed for 15 min. in 2×SSC, 0.1% SDS; 15 min. 0.5×SSC, 0.1% SDS; 15 min. 0.2×SSC, 0.1% SDS at room temperature, and 30 min. in 0.2×SSC, 1% SDS at 50°C.

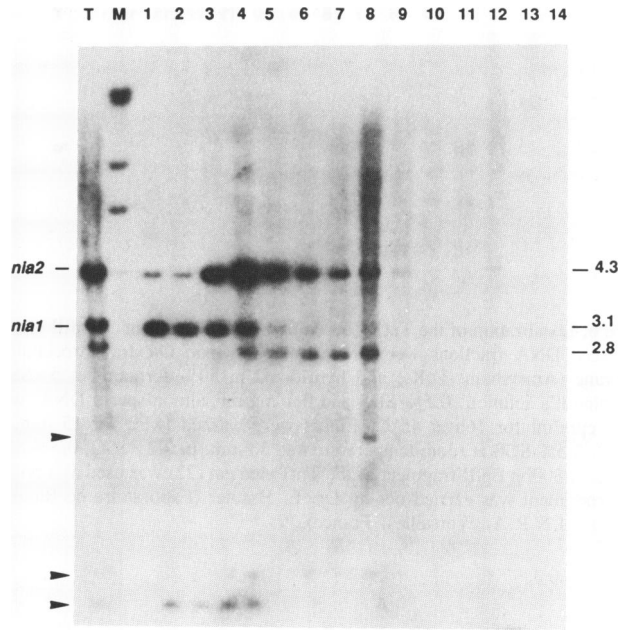

Figure 4. Localization of *nia1* and *nia2* nitrate reductase genes. The probe was a 1.6 kb EcoRI tobacco nitrate reductase cDNA fragment inserted into pBM 102010 plasmid (19). Hybridization was performed in 5×SSPE, 5×Denhardt's solution, 0.5% SDS, 0.1 mg/ml salmon sperm DNA with radiolabeled probe (1.2×10^6 cpm/ml) at 65°C. Filters were washed twice 20 min. in 2×SSPE, 0.5% SDS at room temperature and 10 min. in 2×SSPE, 0.5% SDS at 65°C. M indicates the molecular weight marker (λ DNA digested with HindIII).

size (arrow); this set was centered on fraction 8. The second set was formed by one main band, 3.1 kb in size, and three faint bands, 2.9, 2.6 and 2.1 kb in size; this set was centered on fraction 4. The apparent bimodality of the 3.1 kb fragment is due to the slight overloading of fraction 8.

(iv) Preliminary results concerning the localization of the rubisco small subunit multigene family indicated the presence of at least 10 hybridization bands in EcoRI digests of total DNA (as in the previous work of Jamet et al., 18), which were shown, on fractionated DNA, to belong to three sets of bands centered on fractions 1, 4 and 8, respectively (not shown).

(v) The hybridization results concerning the nitrate reductase genes (Fig. 4) showed three main bands (2.8, 3.1 and 4.3 kb) on unfractionated DNA, in essential agreement with Vaucheret et al. (19). These authors identified the 3.1 kb and the 4.3 kb bands with *nia1* and *nia2*, respectively; *nia1* and *nia2* are the two nitrate reductase genes present in tobacco. When hybridization was performed on DNA fractions, bands corresponding to the *nia1* and *nia2* genes were centered on fraction 1 (1.6949 g/cm^3) and fraction 4 (1.6956 g/cm^3), respectively. The other hybridization band (2.8 kb), due to a *nia*-related sequence (on the basis of the sequence homology revealed by hybridization) not yet investigated, showed a pattern centered on fraction 8. In addition, three weaker bands (arrows) were detected at lower molecular weights, each of them sharing the distribution of one of the bands mentioned before.

(vi) The hybridization pattern obtained with the tobacco β-1,3-glucanase probe (20) is displayed in Fig. 5. Unfractionated DNA showed four hybridizing bands (the two lower molecular weight ones are not visible in this experiment), in agreement with previous results obtained on another cultivar (21). The 14.5 kb band was centered on fraction 8, whereas the 9.6 kb band on fraction 7 and extending into two GC-poorer fractions. The two

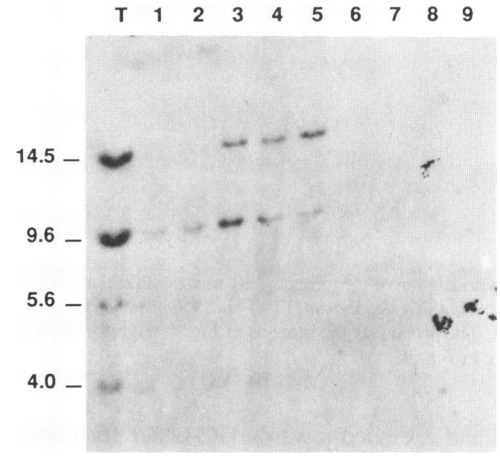

Figure 5. Localization of the glucanase gene. The probe was the PstI insert of the tobacco β-1,3-glucanase cDNA clone pGL43 (21). Hybridization and washing conditions as in Fig. 4.

weaker bands (4.0 and 5.6 kb) could not be assigned with certainty to compositional fractions.

Localization of the multicopy Tnt1 transposable element in tobacco DNA fractions

Two main hybridization bands, 3.8 and 4.7 kb in size, were produced using as a probe a specific Tnt1 sequence (22) on DNA fractions digested with HindIII (Fig. 6). The two bands were localized in a GC-poor region of the genome, but exhibited different distributions in that the 3.8 kb band was centered on

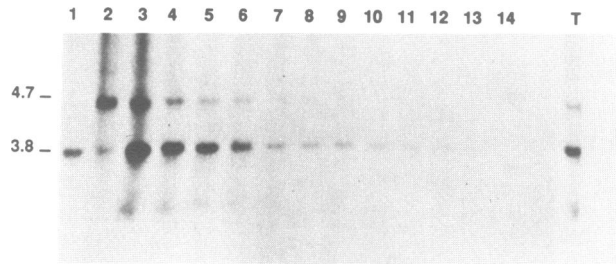

Figure 6. Localization of the Tnt1 tobacco transposable element. HindIII digests of tobacco DNA fractions were blotted onto Hybond C-extra nitrocellulose membrane (Amersham, U.K.) and hybridized in 50% formamide, 5×SSC, 5×Denhardt's solution, 0.5% SDS and 0.1 mg/ml salmon sperm DNA, with 5×10^6 cpm/ml for 16h at 45°C. Filters were washed twice for 15 min. in 2×SSC, 0.5% SDS at room temperature and 30 min. in 0.2×SSC, 0.5% SDS at 65°C. A 300bp BglII fragment of the Tnt1 element (22) was used as a probe. This experiment was carried out by Dr. E. Huttner (Laboratoire de Biologie Cellulaire, I.N.R.A., Versailles, France).

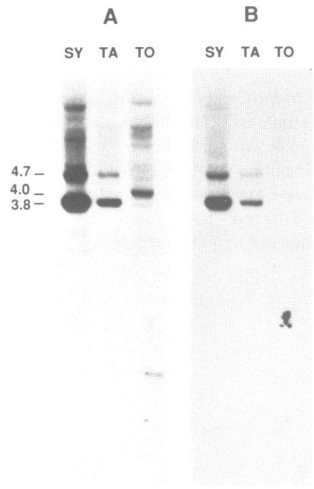

Figure 7. Hybridization of the tobacco Tnt1 transposable element on total DNAs from *N.sylvestris* (SY), *N.tabacum* (TA) and *N.tomentosiformis* (TO) digested with HindIII. The probe was the same as in Fig. 6. Hybridization and washing conditions as in Fig. 4.

fraction 3 and extended towards GC-richer fractions, whereas the 4.7 kb band was essentially present in fractions 2 and 3. This suggests the existence of at least two different sub-families of Tnt1 elements, which comprise over 100 copies (22). This suggestion is supported by the fact that (i) the 3.8 kb band corresponds to an internal HindIII fragment of the sequence, as deduced from the restriction map (22); and (ii) that the 4.7 kb band is also likely to correspond to an internal fragment due to the loss of a HindIII site comprised between the 3.8 kb fragment and a contiguous 0.9 kb fragment. Several faint bands could also be detected.

Hybridization of the Tnt1 probe on total DNAs from tobacco and from its two parental species (Fig. 7A) revealed complex band patterns, as expected. Two strong bands (3.8 and 4.7 kb) were seen in *N.sylvestris* and one major band (4.0 kb) in *N.tomentosiformis*. Tobacco DNA showed the two strong bands derived from *N.sylvestris* and a very weak one corresponding

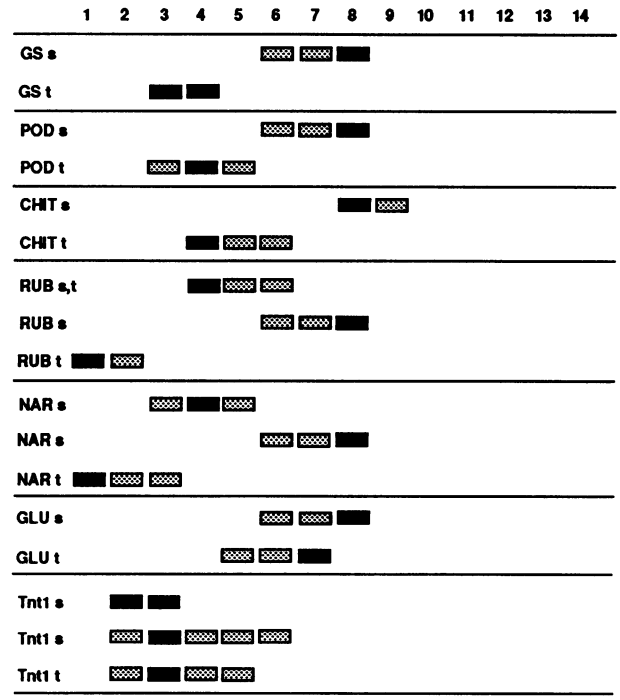

Figure 8. Scheme of the gene localization in the compositional fractions of tobacco DNA. The distribution in the fractions is indicated by black and grey boxes, corresponding to higher and lower hybridization intensities, respectively. Genes are indicated by their acronyms (see Table 2), s and t indicating their origin from *N.sylvestris* and *N.tomentosiformis*, respectively.

to the major band of *N.tomentosiformis*. This *N.tomentosiformis* band could also be detected in several experiments on fractionated tobacco DNA (although this is not visible in Fig. 6), where it seemed to exhibit the distribution of the 3.8 kb band.

The difference between the intensity of hybridization of the two parental species could be due to a different number of copies of the transposable sequence, to a nucleotide divergence, or both. Hybridization at higher stringency led to the disappearance in DNAs from *N.tabacum* and *N.tomentosiformis* of bands originating from *N.tomentosiformis*, but not of the two bands from *N.sylvestris* (Fig. 7B). This result could be expected, since the probe was a Tnt 1 sequence originally isolated from *nia2* (in which it was inserted; 22), a gene which was derived from *N.sylvestris* (see Discussion). The high stringency hybridization indicates that a remarkable nucleotide divergence exists between the transposable elements from the two parental genomes. Whether there is also a difference in the copy number of the transposable elements present in the two parental genomes remains to be established. If such were the case, the copy numbers contributed to the tobacco genome by the two parental genomes might be different. Finally, one should mention that the possibility of a transfer of Tnt 1 elements from one parental genome to the other in the tobacco nucleus cannot be altogether ruled out.

DISCUSSION

Gene distribution is bimodal in the tobacco genome

The hybridization results are summarized in Fig. 8. The most striking feature is that band patterns, as obtained with gene probes, essentially correspond to two sets of DNA fractions having different GC levels.

Indeed, in five cases out of six the distribution of genes or small multigene families in tobacco DNA fractions was clearly bimodal. (i) In three cases (glutamine synthase, peroxidase and endochitinase), the first set was centered on fractions 4 ($\varrho = 1.6956$ g/cm^3), the second one on fraction 8 ($\varrho = 1.6967$). (ii) In two other cases, those of rubisco and nitrate reductase, the distribution was still bimodal but shifted towards fractions of lower buoyant densities (fractions 1 and 4, respectively), whereas additional bands (like the 2.8 kb band of nitrate reductase related sequence, see Fig. 4), showed a distribution centered on fraction 8 (see below). It should be stressed that the bimodal distribution just described was characterized by a buoyant density interval which was comparable to that found for the other three genes mentioned above.

In the case of the glucanase gene (Fig. 5), the results were different from those just described in that the two sets of bands were centered on two closer fractions, 7 and 8, respectively. The results were also expectedly different for the Tnt1 multicopy transposable elements (Fig. 6).

In the following sections, we will first discuss the bimodality of the distribution of the first five genes, and then the cases of the glucanase gene and of the Tnt 1 family.

The bimodal gene distribution in the tobacco genome is due to its amphidiploid nature

An explanation for the bimodal distribution of hybridization patterns obtained with the gene probes can be thought of, if one considers that tobacco is an amphidiploid species (2n=48) derived from *N.sylvestris* (2n=24) and *N.tomentosiformis* (2n=24), which are the female and male parent, respectively (23−25). The two distributions might then correspond to those present in each parental species. In fact, the chromosomes of the latter remain separate in tobacco, due to non-pairing at meiosis (26), even though low rates of homologous pairing of metaphase chromosomes (23,27) and intergenomic exchange of chromosomal segments (28), were reported. This explanation was tested by establishing the parental origin of the hybridization bands. Five of the probes utilized in this work were tested by hybridization on total DNA from *N.tabacum* (cv. Xanthi XHFD8), *N.sylvestris* and *N.tomentosiformis* digested with EcoRI, except in the case of Tnt1, where HindIII digestion was used.

The results obtained, as well as similar data from the literature, are presented in Table 2 and Fig. 8. In the case of the rubisco multigene family (not included in Table 2), bands centered on fraction 1 could be assigned to *N.tomentosiformis*, those centered on fraction 8 to *N.sylvestris*, and those centered on fraction 4 to both parental genomes. In the case of the nitrate reductase gene, the *N.tomentosiformis*-derived band corresponding to *N.sylvestris*-derived band located on fraction 8 was missing or could not be detected.

If the parental origin of the genes is taken into account, the results summarized in the preceding section show that of all genes just discussed, those originated from *N.tomentosiformis* were found in GC-poorer DNA fragments as compared with the related (orthologous or paralogous) sequences from *N.sylvestris*. In other words, we were able to establish that the two parental genomes are compositionally distinct. Strictly speaking this was demonstrated for the isochores hosting the genes tested, as well as for the related sequences corresponding to about 30 additional hybridization bands. However, the number of loci explored and the fact that they were chosen at random, strongly indicate that the compositional differences concern the whole genomes.

Table 2. Parental origin of hybridization bands in tobacco.

Gene (a)	Hybridization band (kb)	Parental origin (b)	References (c)
GS	11.2	T	(d)
	7.1	S	
POD	9.0	T	(16;e)
	7.5	S	
	6.5	S	
	4.4	S	
CHIT	5.7	S	(36;e)
	3.1	T	
	2.9	T	
	2.6	T	
	2.1	T	
	<0.5	S	
Rubisco	see text	see text	(18;e)
NAR	4.3	T	(19)
	3.1	S	
	2.8	S	
GLU	14.5	S	(21;e)
	9.6	T	
	5.6	S	
	4.0	T	
Tnt1	4.7	S	(22;e)
	4.0	T	
	3.8	S	

(a) Abbreviations: GS=chloroplastic glutamine synthase; POD=lignin-forming peroxidase; CHIT=endochitinase; NAR=nitrate reductase; GLU=β-1,3-glucanase.
(b) S=*N.sylvestris* ; T=*N.tomentosiformis*.
(c) If two references are given, the first one is a general reference and the second a reference concerning the assignment of hybridization bands.
(d) B. Hirel, personal communication
(e) Present work.

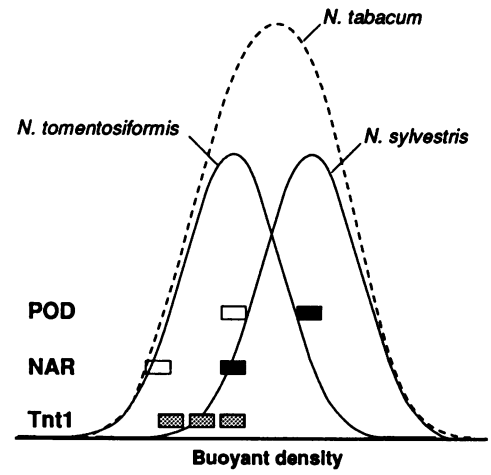

Figure 9. Scheme of the DNA and gene distribution in the amphidiploid genome of tobacco. The two gaussian curves (shifted apart from each other for the sake of clarity) represent the DNA distributions of *N.tomentosiformis* and *N.sylvestris*, respectively; the envelope of the two curves (broken line) represents that of *N.tabacum*. Black boxes represent genes (NAR=nitrate reductase; POD=lignin forming peroxidase) localized in the GC-rich and GC-poor isochores of *N.sylvestris*; white boxes related genes localized in the GC-rich and GC-poor isochores of *N.tomentosiformis*. Grey boxes represent the distribution of Tnt1.

This situation is visualized in Fig. 9, which presents a scheme of the bimodal distribution of genes superimposed on a tentative compositional distribution of DNA fragments, as deduced from the hybridization results, of the tobacco genome and those of its

parental genomes. In the case of the lignin forming peroxidase (which is also the case of glutamine synthase and chitinase), the two genes were localized in the GC-rich fractions of each genome. In the case of the nitrate reductase (and rubisco) genes, hybridization bands are shifted towards GC-poorer fractions in both genomes.

Noteworthy is the fact that no hybridization bands were detected in the GC-richest fractions of the tobacco genome, which are the most heavily methylated (Matassi G., Melis R., Kuo K.C., Gehrke C.W. and Bernardi G., paper in preparation). This finding (similar to that previously reported for the genomes of *Gramineae*; 5) suggests that the GC-richest fractions of tobacco are likely to correspond at least largely to repeated sequences (as also indicated by the presence of bands after digestion with EcoRI and HindIII); incidentally, ribosomal DNA was found, by hybridization with an appropriate probe, in fractions 13 and 14.

The cases of the glucanase genes and of the Tnt1 transposable element

We will now discuss two apparent exceptions to the rule of the bimodal distribution of genes in the tobacco genome. The reasons for such exceptions are very different in the two cases.

The results obtained with the glucanase gene probe is different from those obtained for other genes. Indeed, while the *N.sylvestris* derived gene exhibits the expected distribution centered on fraction 8, already found for other genes derived from *N.sylvestris*, the *N.tomentosiformis* derived gene is centered on a close fraction, 7, with a tailing towards GC-poorer fractions. In this case, the obvious conclusion is that the genomic environment of the *N.tomentosiformis*-derived glucanase gene is different from those found for all other genes derived from *N.tomentosiformis*.

As far as Tnt1 is concerned, the distributions of the *N.sylvestris*-derived fragments were centered on tobacco DNA fractions which were GC-poorer than for the *N.sylvestris*-derived genes. In fact, the 4.7 kb band probably occupies the GC-poorest fractions of *N.sylvestris* (unless there is a movement of transposable elements between the two parental genomes, in which case it could also occupy fractions from *N.tomentosiformis* DNA). The distribution of the *N.tomentosiformis*-derived (4.0 kb) band appeared to show in several experiments the same distribution as the 3.8 kb band of *N.sylvestris*. These findings indicate that there is no bimodality for the *N.sylvestris*-derived Tnt 1 family (and its two sub-families) and the cross-hybridizing elements from *N.tomentosiformis* (see Figs. 8 and 9). This conclusion is not surprising in view of the copy number, the mobility and the divergence of the sequences (see Fig. 7).

The concentration of the Tnt1 family of transposable elements in the GC-poor fractions of the tobacco genome is reminiscent of the compositionally compartmentalized integration of mobile and viral sequences in the genomes of warm-blooded vertebrates (29).

The implications of the bimodality of gene distribution in the tobacco genome

Absence of recombination and nuclear architecture. The findings just discussed support the classical notion (25) that the two parental genomes do not undergo any noticeable recombination process in tobacco. Strictly speaking, our data show no recombinant DNA segments carrying the genes tested and having a size of 100−200 kb. By themselves, they do not rule out exchanges at higher size levels (which are, however, not supported by other notions), nor exchanges involving very small size segments. The only exception could be that of the glucanase gene from *N.tomentosiformis*, although other explanations are possible (see below). In this connection, it should be mentioned that Sperisen et al. (30) have recently demonstrated the possibility of rare events of intergenomic exchange of DNA in the evolution of the tobacco glucanase gene family.

The lack of recombination seen in the present work fits with the concept of *nuclear architecture*, namely with the concept of spatial ordering of chromosomes such as that seen in sexual hybrids between barley and rye (see 31, for a review). Interestingly, a spatial chromosome distribution was also reported in the case of hybrids between *N.plumbaginifolia* and *N.sylvestris*. Telocentric chromosomes of *N.plumbaginifolia* were positioned predominantly at the periphery of metaphase plates, whereas *N.sylvestris* chromosomes occupied the center (32).

The origin of the different composition of parental genomes. The difference in GC levels between the two parental species raises the question of its origin. Two possibilities should be considered here : (a) The amplification of a repeated interspersed sequence changing the composition of one genome and increasing its size (genome sizes for *N.sylvestris*, *N.tomentosiformis* and *N.tabacum* were reported to be 4.2, 3.7 and 7.8 pg per haploid genome, respectively, 33). (b) A change in the overall base composition due to the fixation of directional mutations (7, 34). This second possibility could be tested if a large number of related coding sequences from the parental species were available, because they would show such changes. Unfortunately, only three pairs of coding sequences are presently available, namely the nitrate reductase, the acetolactate synthase (35) and the glucanase genes.

In any case, the amplification of interspersed repeated sequences appears to be unlikely as an explanation for the difference in GC levels between the parental species because the bimodality of gene hybridization patterns indicates that the base composition in intergenic sequences flanking the genes tested is characterized by a constant difference. This should imply a very particular distribution of the hypothetical interspersed repeats. We consider, therefore, that the second explanation is more likely to be correct, the difference in DNA amount being rather due to changes in the (tandemly) repeated sequences which are apparently present at the GC-rich end of the compositional distribution of DNA fragments.

The evolutionary conservation of the parental genomes. The results presented here indicate a very high degree of evolutionary conservation both of isochore composition within each genome, and of gene localization in isochores. In fact, if isochores had changed base composition, or if genes had moved to other compositional compartments, one would not find the same, consistent difference in the GC levels of isochores harboring related (orthologous or paralogous) genes. This conclusion is in agreement with that drawn by Okamuro and Goldberg (10) on the basis of DNA-DNA reassociation kinetic studies.

At least in one case, that of the glucanase gene, the *N.tomentosiformis*-derived copy might have moved to isochores which have an increased GC level compared to the isochores hosting all other genes having the same origin. Alternatively, the *N.tomentosiformis* isochore containing the glucanase gene might have undergone a compositional change.

ACKNOWLEDGEMENTS

We thank Dr. M.Caboche for his interest and support, J.Fleck, B.Hirel, E.Huttner, R.Lotstein (Ciba-Geigy Corp.), F.Meins Jr., for the gift of DNA probes. We also thank Dr. E.Jamet for the gift of total DNA from *N.sylvestris* and *N.tomentosiformis*, and Dr. E.Huttner for his help in some experiments. We thank the French Ministry of Foreign Affairs for a fellowship to G.M. R.M. is indebted to INRA (*Institut National pour la Recherche Agronomique*) for a fellowhip. G.Macaya thanks the *Commission for the European Communities* for financial support. We are grateful to Dr. F.Meins Jr. for critical reading of the manuscript.

REFERENCES

1. Bernardi, G., Olofsson, B., Filipski, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotival, M. and Rodier, F. (1985) *Science*, **228**, 953–958.
2. Bernardi, G. (1989) *Ann. Rev. Genet.*, **23**, 637–661.
3. Salinas, J., Matassi, G., Montero, L.M. and Bernardi, G. (1988) *Nucleic Acids Res.*, **16**, 4269–4285.
4. Matassi, G., Montero, L.M., Salinas, J. and Bernardi, G. (1989) *Nucleic Acids Res.*, **17**, 5273–5290.
5. Montero, L.M., Salinas, J., Matassi, G. and Bernardi, G. (1990) *Nucleic Acids Res.*, **18**, 1859–1867.
6. Bernardi, G. and Bernardi, G. (1986) *J. Mol. Evol.* **22**, 363–365.
7. Bernardi, G., Mouchiroud, D., Gautier, C. and Bernardi, G. (1988) *J. Mol. Evol.* **28**, 7–18.
8. Aissani B., D'Onofrio G., Mouchiroud D., Gardiner K., Gautier C. and Bernardi G. (1991) *J. Mol. Evol.* **32**, 497–503.
9. Mouchiroud D., D'Onofrio G., Aïssani B., Macaya G., Gautier C. and Bernardi G. (1991) *Gene* **100**, 181–187.
10. Okamuro, J.K. and Goldberg, R.B. (1985) *Mol. Gen. Genet.*, **198**, 290–298.
11. Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) *Pl. Molec. Biol. Reporter*, **1**, 19–21.
12. Cortadas, J., Macaya, G. and Bernardi, G. (1977) *Eur. J. Biochem.*, **76**, 13–19.
13. Kirk, J.T.O. (1967) *J. Mol. Biol.*, **28**, 171–172.
14. Szybalski, W. and Szybalski, E. (1971) : In : Cantoni, G.L. and Davidson, D.R. (Eds.) *Procedures in Nucleic Acid Research*, vol. **2**, pp. 311–354.
15. Kemp, J.D. and Sutton, D.W. (1976) *Biochim. Biophys. Acta*, **425**, 148–156.
16. Lagrimini, L.M., Burkhart, W., Moyer, M. and Rothstein, S. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 7542–7546.
17. Shinshi, H., Mohnen, D. and Meins, F. Jr. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 89–93.
18. Jamet, E., Durr, A. and Fleck, J. (1987) *Gene*, **59**, 213–221.
19. Vaucheret, H., Vincentz, M., Kronenberger, J., Caboche, M. and Rouzé, P. (1989) *Mol. Gen. Genet.*, **216**, 10–15.
20. Mohnen, D., Shinshi, H., Felix, G. and Meins, F. Jr. (1985) *EMBO J.*, **4**, 1631–1635.
21. Shinshi, H., Wenzler, H., Neuhaus, J.-M., Felix, G., Hofsteenge, J. and Meins, F., Jr. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 5541–5545.
22. Grandbastien, M.-A., Spielmann, A. and Caboche, M. (1989) *Nature*, **337**, 376–380.
23. Goodspeed, T.H. (1954) The Genus Nicotiana. *Chronica Botanica Co.*, Waltham, Mass.
24. Sheen, S.J. (1972) *Evolution*, **26**, 143–154.
25. Gray, J.C., Kung, S.D., Wildman, S.G., Sheen, S.J. (1974) *Nature*, **252**, 226–227.
26. Stebbins, G.L. (1950) *Variation and evolution in plants*. Columbia University Press, New York.
27. Gerstel, D.U. (1966) *Hereditas* Suppl. **2**, 481–504.
28. Cameron, D.R. (1952) *Genetics*, **37**, 288–296.
29. Rynditch, A., Kadi, F., Geryk, J., Svoboda, J. and Bernardi, G. (1991) *Gene*, (in press).
30. Sperisen, C., Ryals, J. and Meins, F., Jr. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1820–1824.
31. Heslop-Harrison, J.S. and Bennett, M.D. (1990) *Trends Genet.* **6**, 401–405.
32. Gleba, Y.Y., Parokonny, A., Kotov, V., Negrutiu, I. and Momot, V. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 3709–3713.
33. Bennett, M.D. and Smith, J.B. (1976) *Philos. Trans. R. Soc. London*, **274**, 227–274.
34. Bernardi, G. and Bernardi, G. (1990) J. Mol. Evol., **31**, 282–293.
35. Lee, K.L., Townsend, J., Tepperman, J., Black, M., Chui, C.F., Mazur, B., Dunsmuir, P. and Bedbrook, J. (1988) *EMBO J.*, **7**, 1241–1248.
36. Shinshi, H., Neuhaus, J.-M., Ryals, J. and Meins, F., Jr. (1990) *Plant Mol. Biol.*, **14**, 357–368.