

Evidence that a 1.6 kilobase region of *Neurospora* mtDNA was derived by insertion of part of the LaBelle mitochondrial plasmid

Frank E.Nargang, Suchira Pande*, John C.Kennell¹, Robert A.Akins^{1,+} and Alan M.Lambowitz¹
Department of Genetics, University of Alberta, Edmonton, Alberta T6G 2E9, Canada and
¹Departments of Molecular Genetics and Biochemistry, and the Biotechnology Center, The Ohio State University, 484 W.12th Avenue, Columbus, OH 43210, USA

Received November 13, 1991; Revised and Accepted January 29, 1992

GenBank accession no. M76986

ABSTRACT

The LaBelle mitochondrial plasmid hybridizes to a small region of the mtDNA of different *Neurospora* species. Here, we show that the region of homology encompasses 1385 bp of plasmid sequence and 1649 bp of mtDNA sequence. Several findings—that the region of homology is not found in the mtDNAs of other organisms, that it includes the C-terminus of the ORF encoding the plasmid DNA polymerase, and that the ORF sequence in the mtDNA is interrupted by insertions—suggest that the region was part of the plasmid that integrated into mtDNA prior to the divergence of *Neurospora* species. Since the LaBelle plasmid has been found in only one *Neurospora* strain, we infer that the plasmid was lost subsequently from most strains. The LaBelle plasmid is transcribed by the host *Neurospora* mitochondrial RNA polymerase and the major promoter is located upstream of the long ORF, within the region of homology to mtDNA. A promoter used for the transcription of the mitochondrial small rRNA is found at a corresponding position in *Neurospora* mtDNA and may have been acquired via integration of the plasmid sequence. Our results provide evidence that an autonomous infectious element may contribute to sequences that functionally constitute an organism's mtDNA.

INTRODUCTION

Mitochondrial plasmids, which have been found in fungi and plants, include a variety of DNA species that are not homologous to the standard mtDNA (1,2,3). In *Neurospora*, mitochondrial plasmids are not found in standard laboratory strains, but are common in strains isolated from nature (1,4). These plasmids include both circular and linear DNA species. Based on DNA-DNA hybridization experiments, most of the circular plasmids have been placed into one of three homology groups—

Mauriceville, LaBelle, and Fiji—named after the geographic location of the initial isolate from each group (5,6,7). A fourth homology group contains small plasmids (VS DNAs) that coexist in the mitochondria of certain strains of *Neurospora* harboring plasmids of the Mauriceville group (8).

The characteristics of many mitochondrial plasmids suggest that they are autonomous infectious elements. Some of the *Neurospora* mitochondrial plasmids have been shown to be transmitted horizontally (9). The Mauriceville and closely related Varkud plasmids of *Neurospora* contain a 710 amino acid open reading frame (ORF) that encodes an 81 kilodalton protein having reverse transcriptase activity (10–13). These plasmids give rise to unit length transcripts and are believed to replicate via an RNA intermediate and reverse transcription step (14). Two linear DNA mitochondrial plasmids, kalilo and maranhar, have been identified in *Neurospora* (15,16) and the kalilo plasmid has been shown to contain two long ORFs that putatively encode DNA and RNA polymerases (17). These and other linear DNA plasmids have terminal inverted repeats with covalently linked terminal proteins and are believed to replicate by a mechanism analogous to linear DNA viruses of similar structure (16,17).

Mitochondrial plasmid sequences have been found integrated in the mtDNA of various organisms. For example, the mitochondria of the S-type male-sterile cytoplasm of maize contain two linear plasmids, S1 and S2, with inverted terminal repeats that recombine with homologous sequences in the normally circular mitochondrial chromosome to give linear chromosomes with plasmid sequences at their ends (18). Fertile revertants have lost both the plasmid versions of S1 and S2 and the linear mitochondrial chromosomes, but have plasmid sequences incorporated into their mtDNA (19). Linear mitochondrial plasmids have also been found in isolates of the fungus *Claviceps purpurea* and in some cases sequences homologous to the plasmids are found integrated into the mitochondrial chromosome (20, 21). In *Neurospora*, integration of mitochondrial plasmid sequences has been associated with

deleterious phenotypic effects. Integration of the kalilo or maranhar plasmids is correlated with premature senescence due to accumulation of suppressive mtDNA molecules containing deletions and rearrangements (15,16). Similarly, growth impaired mutants of strains carrying plasmids of the Mauriceville homology group contain defective mtDNAs into which plasmid sequences have integrated (22,23).

We have recently reported the complete DNA sequence of the LaBelle mitochondrial plasmid, a 4.1 kb circular DNA found in the LaBelle-1b strain of *Neurospora intermedia* (24). A map of the LaBelle plasmid is shown in Fig. 1. Although unrelated to the Mauriceville group plasmids, the LaBelle plasmid was also found to give rise to transcripts of approximately unit length containing the information for a long ORF of 1151 amino acids. Biochemical studies showed that the LaBelle plasmid ORF encodes a DNA polymerase protein of 120 kDa, which presumably functions in replication of the plasmid (25). Thus, the LaBelle plasmid also appears to be an autonomous infectious element that replicates in *Neurospora* mitochondria.

Of the *Neurospora* mitochondrial plasmids studied thus far, only the LaBelle plasmid hybridizes to wild-type *Neurospora* mtDNA (6). In the present report, we describe the region of similarity between the plasmid sequence and the mtDNA of both *N. crassa* strain 74A and *N. intermedia* strain LaBelle. Our results suggest that the sequence arose by the integration of part of the plasmid into mtDNA prior to the divergence of *Neurospora* species and that the plasmid was subsequently lost from most strains. Further, the integrated sequence includes a promoter that is now used for transcription of the *Neurospora* mitochondrial small rRNA, as well as a possible mtDNA origin of replication. These results indicate that an autonomous infectious element can contribute sequences that functionally constitute an organism's mtDNA.

MATERIALS AND METHODS

Strains of *Neurospora* and growth conditions

The strains used in this study were *N. intermedia* LaBelle-1b (FGSC # 1940), the standard *N. crassa* laboratory strain 74A, and [*poky*] (13–6), an extranuclear mutant deficient in mitochondrial protein synthesis (26). The [*poky*] mutant appears to accumulate mitochondrial RNA polymerase activity and was used to make mitochondrial lysates for *in vitro* transcription assays (27).

Isolation of mtDNA, mitochondrial plasmid DNA, and construction of recombinant plasmids

mtDNAs and mitochondrial plasmid DNAs were isolated as described (28,29). The DNAs were digested with various restriction endonucleases and cloned into bacterial vectors for DNA sequencing and analysis of mtDNA promoters. The following recombinant plasmids contain restriction fragments of mtDNA: pHBE6, *EcoRI*-6 from *N. crassa* 74A mtDNA in pBR322; pNK1, *EcoRI*-6 from *N. intermedia* LaBelle-1b mtDNA in pUC19; pLK8, 1170 bp *KpnI*-8 fragment from LaBelle mtDNA in Bluescribe vector pBS(+)(Stratagene, La Jolla, CA); pKK8–5, *KpnI*-8 fragment from 74A mtDNA in pBS(+). The following recombinant plasmids contain restriction fragments of LaBelle plasmid DNA. pLSLE38, entire LaBelle plasmid linearized at the *EcoRI* site and cloned in pBR322; pLSP1, entire LaBelle plasmid linearized at the *PstI* site and cloned in pBR322 (28); p1940LB-14, entire LaBelle plasmid linearized at the *BamHI* site and cloned in the *BamHI* site of pBS(+); pLBE, 556

bp *BamHI*-*EcoRI* fragment of the LaBelle plasmid (positions 3516 to 1/4070) subcloned from pLSP1 in pBS(+); pLXE, 416 bp *XbaI*-*EcoRI* fragment of the LaBelle plasmid (positions 3656 to 1/4070) subcloned from pLSP1 in pBS(+).

Southern hybridization

To localize regions of homology between the LaBelle plasmid and *Neurospora* mtDNA restriction digests of DNA from LaBelle-1b and wild-type 74A mitochondria were electrophoresed in 1% agarose/ethidium bromide gels. The gels were blotted to nylon membranes (Hybond N, Amersham, Arlington Heights, IL), and the blots were hybridized at low stringency (10% formamide, 37°, 5×SSC), with ³²P-labeled p1940LB-14, which contains the LaBelle plasmid *BamHI* clone. The blots were then washed and autoradiographed.

DNA sequencing

Procedures used for the isolation, manipulation, and sequencing of DNAs were as described previously (24). All sequences were determined completely on both strands. Comparisons to database sequences were performed using the programs of PCGENE (Intelligenetics, Mountain View, CA) with release 25 of the EMBL nucleic acid sequence database and release 17 of the Swiss protein database.

In vitro transcription assays

In vitro transcription reactions were carried out using mitochondrial lysates from the [*poky*] mutant of *N. crassa* or the LaBelle-1b strain of *N. intermedia*, essentially as described (27). DNA templates were prepared for *in vitro* transcription by digestion with restriction enzymes, followed by extraction with phenol-chloroform-isoamyl alcohol (25:24:1) and precipitation with 2.5 volumes of ethanol in the presence of 0.8 M ammonium

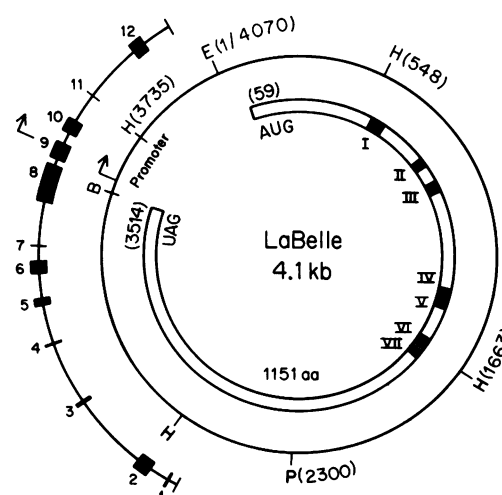


Figure 1. Map of the LaBelle plasmid. The long ORF, encoding a protein with DNA polymerase activity (25), is indicated in the inner ring with its start and stop codons. The darkened areas numbered I through VII in the ORF indicate the regions of amino acid similarity to reverse transcriptases (24). The region of homology to mtDNA is indicated by the arc outside the plasmid. The darkened areas on this arc signify gaps 1 through 12, as described in the text. Arrows indicate location of the major plasmid promoter and the mtDNA promoter in gap 9. The numbers in parentheses are nucleotide positions in the LaBelle plasmid according to Pande *et al.*, (24). Abbreviations: E, *EcoRI*; H, *HincII*; P, *PstI*; B, *BamHI*.

acetate. Restriction enzyme digests and DNA concentrations were checked by agarose gel electrophoresis.

Transcription reactions were carried out in 25 μ l of solution containing 10 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 100 μ g per ml acetylated bovine serum albumin (Gibco BRL, Gaithersburg, MD), 150 μ M ATP, CTP and GTP, and 25 μ M UTP (Pharmacia, Piscataway, NJ), 25 μ Ci [α -³²P]UTP (3,000 Ci/mmol, ICN, Costa Mesa, CA), approximately 500 ng DNA template, and 5 μ g of mitochondrial lysate protein. To suppress nonspecific nuclease activity, mitochondrial lysates were preincubated with 0.5 volumes of an antibody preparation against *N. crassa* endo-exonuclease (30) in 10 mM Tris-HCl, (pH 7.5) for 30 to 60 min at 4°, immediately before use. *In vitro* transcription reactions were initiated by addition of mitochondrial lysate, incubated at 30° for 5 min, and chased with 100 μ M UTP for an additional 5 min. After the chase period, reactions were terminated by adding 25 μ l of solution containing 2 M ammonium acetate, *E. coli* tRNA (250 μ g per ml; Sigma Chemical Co., St. Louis, MO), followed immediately by extraction with phenol-chloroform-isoamyl alcohol (25:24:1) and precipitation with ethanol. To monitor recovery of products, a ³²P-labeled *in vitro* transcript of 117 nt, synthesized from *Kpn*I-linearized pBluescript SK(+) vector using bacteriophage T7 RNA polymerase (27), was added prior to phenol-extraction.

For gel electrophoretic analysis, the ethanol-precipitated RNA products were rinsed with 70% ethanol, dried, and resuspended in 4 μ l of 90% formamide, 90 mM Tris, 90 mM boric acid, 1 mM EDTA. The products were heated to 65° for 3 min and analyzed by electrophoresis in a 6% polyacrylamide gel containing 8 M urea, 90 mM Tris, 90 mM boric acid, and 1 mM EDTA. The gel was dried and autoradiographed. Size markers were *Sau*3AI fragments of pBS(+), which had been 5' end-labeled with [γ -³²P]ATP (New England Nuclear, Boston, MA) using T4 polynucleotide kinase (Gibco BRL, Gaithersburg, MD) and denatured by heating to 95° for 3 min immediately prior to loading on the gel.

The 5' end of an *in vitro* transcript was mapped by primer extension with AMV reverse transcriptase (Gibco BRL, Gaithersburg, MD), using a 5' end labeled synthetic oligonucleotide (5'-CCTGACGTAATGTCATGG-3') complementary to positions 3735–3718, as described (27).

RESULTS

A region of homology between mtDNA and the LaBelle plasmid

We reported previously that the LaBelle mitochondrial plasmid hybridizes weakly to the *Eco*RI-6 fragment of *N. crassa* strain 74A mtDNA (6). Additional Southern hybridization experiments localized the region of homology to two adjoining *Kpn*I fragments (*Kpn*I-8 and -9) extending through the border of *Eco*RI fragments

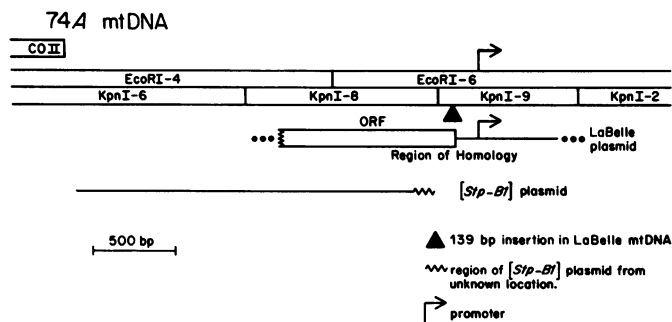
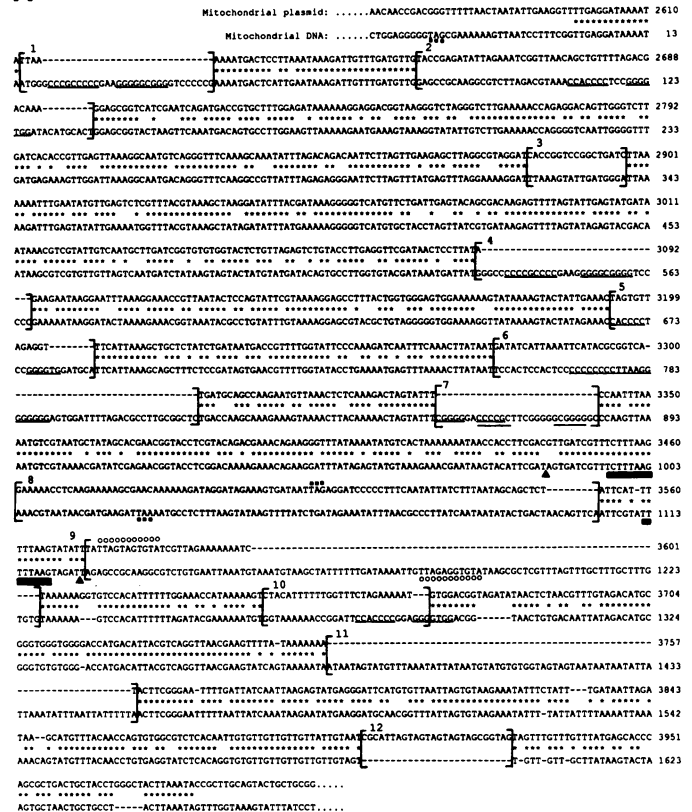


Figure 2. Map of the *Eco*RI-4/-6 region of *N. crassa* 74A mtDNA showing the region of homology to the LaBelle plasmid. The Figure shows the extent of the ORF in the homologous region of the LaBelle plasmid, the location of the 139 bp insertion in *N. intermedia* LaBelle mtDNA, the locations of sequences corresponding to the [*stpB1*] plasmid (36; see Discussion), and the positions of plasmid and mtDNA promoter sequences. COII indicates a portion of the coding region for cytochrome *c* oxidase subunit 2 in *Eco*RI-4.

A



B

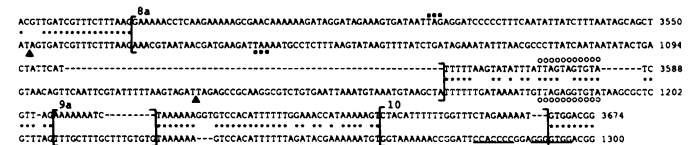


Figure 3. A. Homology between the LaBelle plasmid and the mtDNA from *N. intermedia* strain LaBelle-1b. The top line is the plasmid sequence and the bottom line the mtDNA sequence. Numbering of the plasmid sequence is from Pandé *et al.* (24) and for the mtDNA begins at the start point of homology. Asterisks between the sequences indicate matching nucleotides. Square brackets indicate gaps 1 to 12, as defined in the text. Matching nucleotides in gaps are not indicated by asterisks. Lines below nucleotides in the gaps indicate GC-rich palindromes in the mtDNA sequence. The three closed squares above the plasmid sequence at positions 3512 to 3514 indicate the stop codons of the long ORF in the LaBelle plasmid. The closed squares below the mtDNA sequence indicate the stop codons flanking the 'mtDNA ORF' (see text). Open circles indicate *Neurospora* mtDNA promoter consensus sequences in Gap 9. The triangles under the mtDNA sequence delineate the region of the *N. intermedia* mtDNA sequence, which is not present in *N. crassa* strain 74A mtDNA. The solid bars under mtDNA sequences indicate repeats flanking gap 8 (see text). B. Alternative alignment in the region of gaps 8 and 9. The promoter sequences have been aligned generating a larger gap 8 and a shorter gap 9, designated gap 8a and gap 9a, respectively.

4 and 6 (Fig. 2). This region of *Neurospora* mtDNA is conserved in all natural isolates and laboratory strains of *Neurospora* that have been examined by restriction analysis (31,32). In order to identify the hybridizing sequences, we determined 1813 bp of sequence of the appropriate region of mtDNA from *N. intermedia* LaBelle-1b. In addition, we determined 1252 bp of sequence from the mtDNA of the standard laboratory wild-type *N. crassa* 74A, corresponding to positions 91 to 401 and 699 to 1778 of the LaBelle-1b mtDNA sequence. These sequences are listed in GenBank under accession number M76986.

A portion of the sequence determined from the LaBelle-1b mtDNA is shown in Fig. 3. The sequence begins in restriction fragment *EcoRI*-4, 325 bp from the junction of the *EcoRI*-4 and -6 fragments (Fig. 2). The mtDNA sequence determined from the wild-type 74A strain is identical to that of LaBelle-1b, except for an insertion of 139 bp in the LaBelle-1b strain, encompassing positions 986 to 1124, and a run of 8 G's at position 782 to 789 in LaBelle-1b compared to 9 in 74A (Fig. 3).

Comparison of the *N. intermedia* mtDNA sequence with the previously determined LaBelle plasmid sequence (24) showed

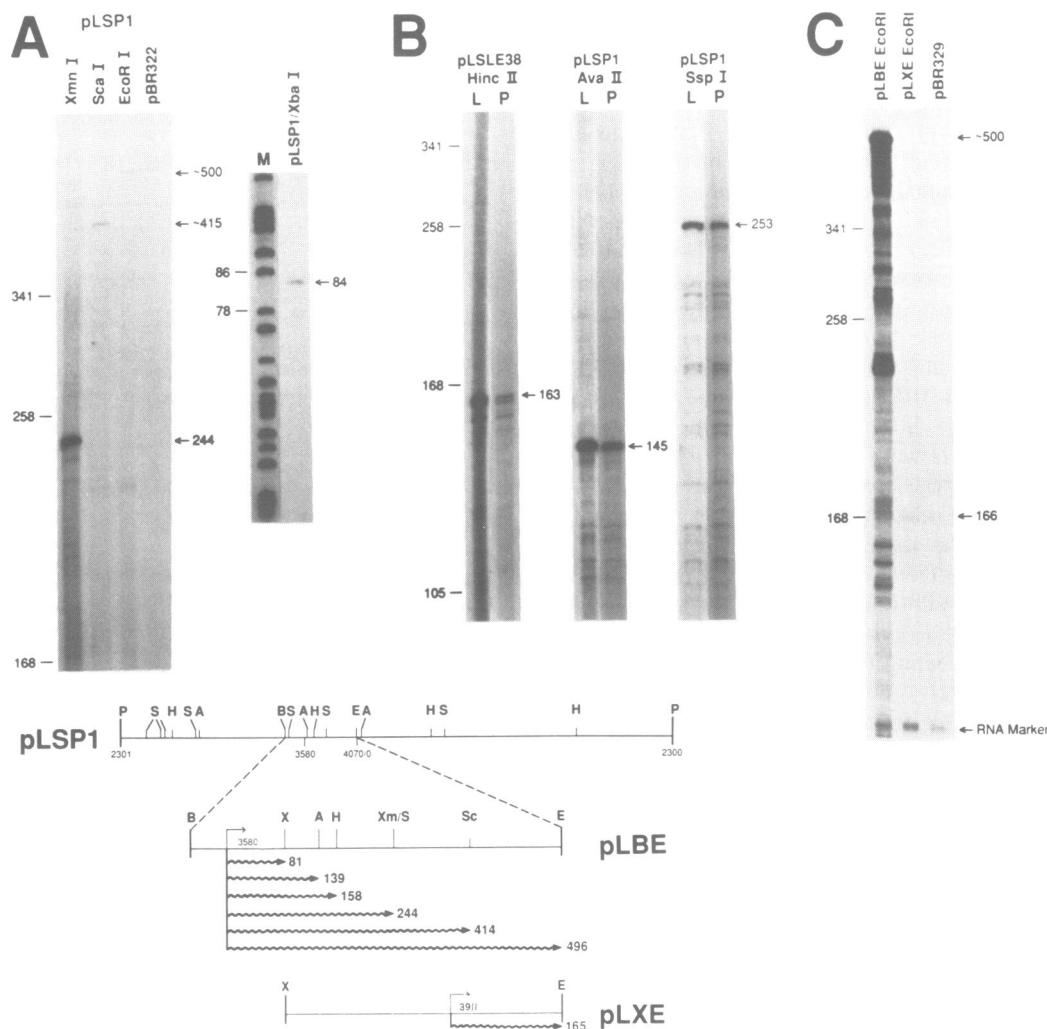


Figure 4. *In vitro* transcription of the LaBelle mitochondrial plasmid. Top panel: A. Identification of *in vitro* transcription initiation sites in the LaBelle plasmid. Templates for *in vitro* transcription reactions were plasmid pLSP1, which contains the entire LaBelle plasmid, digested with *XmnI*, *ScaI*, *EcoRI*, or *XbaI*. *In vitro* transcription was carried out as described in Materials and Methods, using mitochondrial lysates from the [*poky*] mutant, which appears to accumulate mitochondrial RNA polymerase. *EcoRI*-digested pBR322 was used as a control. M, sequencing ladder used to provide size markers. B. Comparison of *in vitro* transcription of the LaBelle plasmid using LaBelle (L) and [*poky*] (P) mitochondrial lysates. Templates were plasmid pLSLE38 digested with *HincII* and pLSP1 digested with *AvaII* or *SspI*. The lanes obtained for the LaBelle lysates required prolonged autoradiogram exposures for bands to have the same intensity as those obtained with [*poky*] lysates. C. Identification of a minor *in vitro* transcription initiation site in the LaBelle plasmid. The regions of the LaBelle plasmid containing the major transcription initiation site and a potential minor transcription initiation site were subcloned separately. The resulting plasmids, pLBE and pLXE, respectively, were digested with *EcoRI* and used as templates for *in vitro* transcription. *EcoRI*-digested pBR329 was used as control. Autoradiograms were exposed for 2 days with an intensifying screen to detect a band due to the minor transcription initiation site. With this time of exposure, the lane for the major initiation site (pLBE) is overexposed, while the pBR329 control lane shows no bands. The multiple bands in the pLBE/*EcoRI* lane are due to premature termination or degradation of the major 500 nt *in vitro* transcript and are pronounced due to overexposure of the autoradiogram. In all three panels, numbers at the left indicate lengths (bp) of size markers (5' end-labeled *Sau3AI* fragments of pBS(+)). Numbers at right of autoradiograms indicate sizes of runoff transcripts calculated from DNA size markers. Bottom panel: Schematic of plasmids and restriction sites used for *in vitro* transcription. The sizes of expected runoff transcripts initiated at position 3580 for the major promoter, or at position 3911 for the minor promoter, are shown below the pLBE and pLXE inserts, respectively. Abbreviations: A, *AvaII*; B, *BamHI*; E, *EcoRI*; H, *HincII*; P, *PstI*; S, *SspI*; Sc, *Scal*; X, *XbaI*; Xm, *XmnI*.

that they could be aligned for a span of 1385 nts of plasmid sequence and 1649 nts of mtDNA sequence (Fig. 3A). To test for additional similarity, we also compared the LaBelle plasmid sequence to unpublished sequence of 74A mtDNA (provided by E. Agsteribbe, State University of Groningen, Groningen, The Netherlands), which extended the region of mtDNA sequence determined in our study by several kb in both directions. No evidence of further homology was found. Comparison of both the plasmid and mtDNA sequences to entries in the organelle portion of the EMBL data base revealed no striking similarities.

The position of the region of homology relative to the LaBelle plasmid and its long ORF is shown in Fig. 1. We define as 'gaps', any region of more than fifteen nts where the sequences cannot be aligned with greater than 40% identity, or where one sequence has an insertion/deletion of more than 6 nts relative to the other. By these criteria, the region of homology contains twelve gaps indicated in Fig. 3A. Gaps 1, 4, 5, 6, and 7 include insertions in the mtDNA that are GC-rich sequences containing palindromic regions (Fig. 3A). Gaps 2, 3, 8, 9, and 10 contain sequences from the plasmid and LaBelle mtDNA that cannot be aligned. The mtDNA sequence in gap 8 corresponds almost exactly to the 139 bp insertion in the LaBelle mtDNA relative to strain 74A. A short direct repeat T(C/T)TTAAG is found near the boundaries of gap 8 in both the plasmid and mtDNA. It is possible that the gap 8 sequence in 74A mtDNA was deleted by recombination between direct repeat sequences or that the direct repeat was generated during insertion of the sequence in the LaBelle strain. The plasmid sequence in gap 10 is 77% AT, while the mtDNA sequence contains a short GC-rich palindrome. Gap 11 contains mtDNA sequence that is 88% AT and has no counterpart in the plasmid. Gap 12 is the only gap where plasmid sequence has no counterpart in the mtDNA. Gap 8 and gap 9 are separated by a short region containing only eighteen matched bases (Fig. 3A). Both the plasmid and mtDNA sequences in gap 9 contain *Neurospora* mtDNA promoter consensus sequences (see below). An alternative alignment based on the similarity of the promoter sequences is shown in Fig. 3B.

Surprisingly, the region of the LaBelle plasmid homologous to mtDNA includes 917 bp at the 3' end of the plasmid's long ORF (Fig. 1). The termination codon of the ORF occurs in gap 8. Prior to gap 8, the only disruptions in the reading frame in mtDNA occur as the result of the insertions or long unaligned sequences corresponding to gaps 1 through 7. All other changes are simple mismatches, including 73 transitions and 91 transversions. By contrast, in the 545 bp of mtDNA sequence following gap 8, the sequences differ by 12 small insertions/deletions, as well as 24 transitions, 22 transversions, and the larger insertions/deletions corresponding to gaps 8 through 12.

When the mtDNA sequence was translated with gaps included, six ORFs of over 75 codons were revealed. None of these had striking similarity to protein sequences in the database. The longest ORF contains 115 codons and occurs at position -22 to position 323 of the mtDNA sequence in Fig. 3. An in-frame ATG codon occurs 12 codons after the start of this ORF. None of the other five ORFs contains an ATG codon in a position that would begin a protein of over 35 amino acids.

By excluding gaps 1 to 7 plus one or two nucleotides flanking the gaps from the LaBelle mtDNA sequence, we derived an ORF sequence, referred to as 'mtDNA ORF', which is related to the C-terminus of the LaBelle plasmid ORF. Comparison of codons in the mtDNA ORF with those in the plasmid shows that 26

codons are altered in the first position, 6 in the second position and 54 in the third position, with 36 codons being affected in more than one position. The 'mtDNA ORF' contains no internal stop codons, but is bounded by stop codons eight codons upstream from the start point of homology with the LaBelle plasmid and, at the end of the ORF, eight codons into gap 8 (Fig. 3A).

The absence of small insertions/deletions in the mtDNA ORF and the high proportion of codon changes occurring in the third position suggest that it was at one time under selective pressure. One possibility is that there was selective pressure to preserve the ORF after it was inserted into mtDNA. An alternate possibility, however, is that because of differences in replication or repair mechanisms, the rate of nucleotide substitutions is slower for sequences in mtDNA than in the mitochondrial plasmid, so that the ORF sequence may have undergone relatively few changes following its insertion into mtDNA, except for the four large insertions corresponding to gaps 1, 4, 6, and 7. In that case, the apparent selective pressure may simply reflect the prior existence of the ORF in the plasmid. This second possibility is favored by comparisons of mtDNAs from different *Neurospora* species and strains, including the comparison of the *EcoRI*-4/6 region in the present work, which show that nucleotide substitutions and small insertions or deletions are rare and that most differences correspond to large insertions or deletions (31,32). At the same time, mitochondrial plasmids harbored in different *Neurospora* strains have a comparatively high rate of nucleotide substitutions relative to their mtDNAs (11,12). If the original ORF sequence is in fact more strongly conserved in the mtDNA than in the plasmid, then the gaps containing relatively

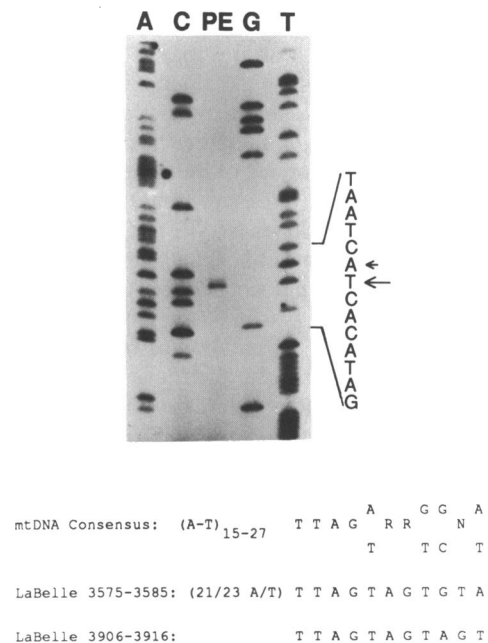


Figure 5. Primer extension mapping of the 5' end of a LaBelle plasmid *in vitro* transcript. Top panel. An *in vitro* transcript was synthesized from pLBE digested with *SspI*, using a [*poky*] mitochondrial lysate, and its 5' end was mapped by primer extension, using a 5' end-labeled synthetic oligonucleotide, as described in Materials and Methods. The primer extension lane (PE) is flanked by sequencing ladders generated from the same 5' end-labeled primer on pLBE DNA template. Bottom panel. Comparison of LaBelle plasmid promoter sequences with the previously determined *Neurospora* mtDNA promoter consensus sequence (27,33).

long unaligned sequences (gaps 2, 3, and 5) would represent sequences that have undergone extensive mutation in the plasmid and are not essential for function of the plasmid ORF.

Identification of promoters

The LaBelle plasmid sequence inserted into mtDNA includes not only the C-terminus of the plasmid ORF, but also a substantial part of the small noncoding region of the plasmid, which is expected to contain elements required for its replication and transcription. To identify transcription initiation sites in the LaBelle plasmid, recombinant plasmids containing the entire LaBelle plasmid or subfragments thereof were digested with different restriction enzymes and used as templates for runoff transcription assays with *Neurospora* mitochondrial lysates. Fig. 4A shows runoff transcripts obtained with templates generated by restriction enzyme digestion of recombinant plasmid pLSP1, which contains the entire LaBelle plasmid cloned in pBR322. Templates consisting of pLSP1 digested with *Xmn*I, *Sca*I, *Eco*RI and *Xba*I gave major runoff transcripts of ca. 244, 415, 500 and 84 nucleotides, respectively. No runoff transcripts were detected with the linearized vector pBR322. Recombinant plasmid pLBE, which contains a subclone of the putative promoter region, gave identical runoff transcripts when digested with *Xmn*I, *Sca*I or *Eco*RI (data not shown). The sizes of the runoff transcripts map the major *in vitro* transcription start site to approximately position 3577 of the LaBelle plasmid sequence (17; Fig. 3).

Inspection of the sequence around position 3577 of the LaBelle plasmid showed that there is a perfect match to the 11 nt *Neurospora* mtDNA promoter consensus (positions 3575–3585), preceded by an AT-rich region (21/23 nts), as found previously for active *Neurospora* mtDNA promoters (Fig. 5, bottom panel; 27,33). Primer extension mapping of the 5' end of a LaBelle plasmid runoff transcript showed a major band corresponding to the A-residue at position 3580, along with a lighter band corresponding to the T-residue one nt upstream (Fig. 5). These bands suggest that the major initiating nucleotides correspond to the sixth and fifth nucleotides of the promoter consensus, as found previously for other *Neurospora* mtDNA promoters (27).

Two full length plasmid clones, pLSP1 and pLSE38, were used to compare runoff transcripts obtained using the standard [*poky*] mitochondrial lysate with a mitochondrial lysate from the LaBelle strain, which contains the plasmid. As shown in Fig. 4B, we detected no differences in the patterns of transcripts, although the LaBelle lysate gave smaller amounts of RNA products necessitating longer autoradiogram exposures.

The promoter at positions 3575–3585 is the only sequence in the LaBelle plasmid that is an exact match to the previously published *Neurospora* mtDNA promoter consensus sequence. Two additional sequences that contain 10/11 matches to the consensus were found at positions 3906–3916 and positions 1680–1690, but these sequences are not preceded by an upstream AT-rich region, which has been found upstream of active promoters in *Neurospora* mtDNA (33). In order to determine if the sequence at 3906–3916 is a weak promoter, which cannot compete effectively for RNA polymerase or transcription factors, we subcloned a 416 bp *Xba*I-*Eco*RI fragment, which separates this sequence from the major promoter at position 3580. Fig. 4C shows that *Eco*RI-digested pLXE, which contains the subcloned region including the potential weak promoter, gave a runoff transcript of the expected size, approximately 166 bases, but this band appeared only after prolonged autoradiogram exposure. The

flanking lanes in Fig. 4C show comparably long exposures of *in vitro* transcripts from the major promoter (pLBE/*Eco*RI) for comparison and for pBR329 cut with *Eco*RI as a control. The sequence at position 1680 may also be a weak promoter, which cannot be detected in clones containing the major promoter, but this was not tested.

Previous studies showed that the *Eco*RI-6 fragment of strain 74A mtDNA contains two active promoters that appear to contribute to synthesis of the mitochondrial small rRNA via precursors that contain 5' end extensions (33). One of these promoters is located within the region of homology (positions 1184–1194) and the other is downstream, 2.5 kb from the *Eco*RI-4/6 junction. The mtDNA promoter at position 1184–1194 and the LaBelle plasmid promoter are located at approximately the same relative position within the region of homology. Surprisingly, however, this position is within gap 9, and the promoters are flanked by sequences that are not highly similar between the plasmid and mtDNA (Fig.4A). An alternative arrangement of the sequence in this region would be alignment of the two promoters as shown in Fig. 3B. This would result in a longer gap 8 (gap 8a) and a shorter gap 9 (gap 9a). The weak plasmid promoter at positions 3706–3916 occurs in gap 12, which has no cognate in the mtDNA sequence.

DISCUSSION

Our results show that a 1.6 kb region of *Neurospora* mtDNA that is conserved throughout the genus (31,32) and that contains a promoter used for transcription of the mitochondrial small rRNA (33), was derived by integration of the LaBelle plasmid. This conclusion is supported by several observations. First, this region is not present in mtDNAs of other organisms sequenced thus far. If the homologous region was originally a functional part of *Neurospora* mtDNA, we anticipate that it would be present in other organisms; otherwise it would have had to have been lost independently from other mtDNAs, which is unlikely. Second, the region of homology includes 305 codons of the C-terminus of the plasmid ORF, which encodes a DNA polymerase that presumably functions in replication of the plasmid (25). Finally, the ORF sequence in *Neurospora* mtDNA is interrupted by several insertions rendering it non-coding in the mtDNA. It seems unlikely that the C-terminal 305 codons of the plasmid DNA polymerase, which is presumably under selective pressure in the plasmid, would have arisen following recombination with a region of the mitochondrial genome that does not contain an ORF. Since the LaBelle plasmid has been found in only a single strain of *N. intermedia*, we infer that the integration occurred prior to the divergence of different *Neurospora* species and that the plasmid was subsequently lost from most strains. Robison *et al.* (34) recently reported a potentially analogous situation involving similarity between a segment of the putative RNA polymerase ORF of a linear mitochondrial plasmid and a mtDNA fragment present in different species of the mushroom *Agaricus*, but they could not decide on the origin of the similar sequences.

The LaBelle plasmid sequence inserted into the mtDNA includes not only the C-terminus of the plasmid ORF, but also a substantial part of the small noncoding region of the plasmid, which is expected to contain elements required for replication and expression of the plasmid. Using an *in vitro* transcription assay, we show that the major promoter is in fact located within this noncoding region, within the region of homology to *Neurospora* mtDNA. The plasmid contains at least one additional

weak promoter, which is also located in the noncoding region, 326 bp downstream of the major promoter. We found no difference in the pattern of *in vitro* transcription between mitochondrial lysates from strains that do or do not contain the LaBelle plasmid. Therefore, our results suggest that the plasmid relies on the host *Neurospora* mitochondrial RNA polymerase and does not encode a component essential for its own transcription. The use of the host RNA polymerase is a common strategy employed by many infectious elements, although to our knowledge this is the first direct evidence that this strategy is used by a mitochondrial plasmid.

The plasmid promoters in the noncoding region are positioned correctly to synthesize transcripts containing the long ORF. Full length transcripts of the LaBelle plasmid have been detected, but they are present at low concentrations, and it has not been possible to map their 5' and 3' ends (24). Given the good match of the major LaBelle plasmid promoter to the *Neurospora* mtDNA promoter consensus sequence, it is somewhat surprising that the LaBelle plasmid transcript is present at such low concentrations. However, since it is not yet clear what features distinguish strong and weak promoters, the small amount of plasmid transcripts could reflect either a slow rate of synthesis or instability of the transcript *in vivo*.

Previous studies showed that the *EcoRI*-6 fragment of *Neurospora* mtDNA contains two functional promoters that appear to contribute to the synthesis of the mitochondrial small rRNA via precursor RNAs that contain 5' end extensions (33). The present results show that one of these promoters is located within the region of homology in a position corresponding to the LaBelle plasmid promoter. We have presented two possible sequence alignments for the region containing these promoters. Either alignment suggests that the sequence flanking the promoters has undergone considerable divergence. In fact, the sequence from the start of gap 8 (or 8a) to the end of gap 9 (or 9a) is the most poorly conserved in this region. In view of the correspondence in the relative positions of the two promoters, it seems likely that the mtDNA promoter was part of the plasmid sequence that integrated into mtDNA and that the surrounding sequences subsequently diverged in the two DNA species.

The *EcoRI*-4/6 region was found previously to be the site of short tandem reiterations in *N. crassa* mutant mtDNAs and was suggested to be the location of a replication origin of *Neurospora* mtDNA (28,35). Recently, Almasan and Mishra (36) described a 2.2 kb plasmid-like DNA that is derived from this region and accumulates in subcultures of the [*stpB1*] mutant. They noted that the mtDNA sequences found in the [*stpB1*] plasmid include a promoter, gyrase cleavage sites, potential secondary structures, and consensus sequences found at mtDNA replication origins in other organisms, and they suggested that these features constitute a replication origin in *Neurospora* mtDNA. Interestingly, the mtDNA sequences contained in the [*stpB1*] plasmid partially overlap the region of homology with the LaBelle plasmid (See Fig. 2), and the region containing the possible mtDNA replication origin corresponds to sequences that were ultimately derived from the plasmid. All of the features identified in the mtDNA sequence by Almasan and Mishra (36) can be identified in the LaBelle plasmid, with varying degrees of similarity, with the exception of the palindromic sequence, which may have been derived from a separate insertion into the mtDNA. However, these features are found in the region corresponding to the C-terminus of the plasmid ORF. Further, the promoter consensus sequence identified by Almasan and Mishra (36), which is at the junction

of *EcoRI*-4/6 in the mtDNA, has diverged from the consensus sequence in the plasmid and is nonfunctional in *in vitro* transcription assays in both the plasmid and mtDNA (see 33), although it could still be functional *in vivo*. If the *EcoRI*-4/6 region contains a mtDNA replication origin derived from the LaBelle plasmid, another replication origin must have been functional prior to insertion of the plasmid sequence. Analysis of defective mtDNAs derived from different regions of *N. crassa* mtDNA do suggest that there are at least two replication origins, one of which may be in the *EcoRI*-4/6 region (37,38).

Finally, our results provide evidence that an autonomous infectious element can contribute to sequences that functionally constitute an organism's mtDNA. Infectious elements may also account for some strain-specific insertions, which are found frequently in fungal mtDNAs (31,32), and for proposed cases of horizontal transmission of group I and group II introns (see 39).

ACKNOWLEDGMENTS

This work was supported by a grant from the NSERC of Canada to F.E.N. and NIH grant GM37949 to A.M.L. S.P. was the recipient of a post-graduate scholarship from the AHFMR.

REFERENCES

- Lambowitz, A.M., Akins, R.A., Garriga, G., Henderson, M., Kubelik, A., and Maloney, K.A. (1985) In Quagliariello, E., Slater, E.C., Palmieri, F., Saccone, C., and Kroon, A.M. (eds.), *Achievements and Perspectives in Mitochondrial Research*, Volume II: Biogenesis. Amsterdam: Elsevier Science Publishers, pp237-247.
- Tudzynski, P., Rogmann, P., and Neuhaus, H. (1986) *Prog. Bot.* **48**, 249-259.
- Lonsdale, D.M., Brears, T., Hodge, T.P., Melville, S.E., and Rottmann, W.H. (1988) *Phil. Trans. R. Soc. Lond. B* **319**, 149-163.
- Nargang, F.E. (1985) *Exper. Mycol.* **9**, 285-293.
- Collins, R.A., Stohl, L.L., Cole, M.D., and Lambowitz, A.M. (1981) *Cell* **24**, 443-452.
- Stohl, L.L., Collins, R.A., Cole, M.D., and Lambowitz, A.M. (1982) *Nucleic Acids Res.* **10**, 1439-1458.
- Natvig, D.O., May, G., and Taylor, J.W. (1984) *J. Bacteriol.* **159**, 288-293.
- Saville, B.J. and Collins, R.A. (1990) *Cell* **61**, 685-696.
- Collins, R.A. and Saville, B.J. (1990) *Nature* **345**, 177-179.
- Nargang, F.E., Bell, J.B., Stohl, L.L., and Lambowitz, A.M. (1984) *Cell* **38**, 441-453.
- Nargang, F.E. (1986) *Mol. Biol. Evol.* **3**, 19-28.
- Akins, R.A., Grant, D.M., Stohl, L.L., Bottorff, D.A., Nargang, F.E., and Lambowitz, A.M. (1988) *J. Mol. Biol.* **204**, 1-25.
- Kuiper, M.T.R., Sabourin, J.R., and Lambowitz, A.M. (1990) *J. Biol. Chem.* **265**, 6936-6943.
- Kuiper, M.T.R. and Lambowitz, A.M. (1988) *Cell* **55**, 693-704.
- Bertrand, H., Chan, B.S.-S., and Griffiths, A.J.F. (1985) *Cell* **41**, 877-884.
- Court, D.A., Griffiths, A.J.F., Kraus, S.R., Russell, P.J., and Bertrand, H. (1991) *Curr. Genet.* **19**, 129-137.
- Chan, B.S.-S., Court, D.A., Vierula, J., and Bertrand, H. (1991) *Curr. Genet.* **20**, 225-237.
- Schardl, C.L., Lonsdale, D.M., Pring, D.R., and Rose, K.R. (1984) *Nature* **310**, 292-296.
- Schardl, C.L., Pring, D.R., and Lonsdale, D.M. (1985) *Cell* **43**, 361-368.
- Tudzynski, P. and Esser, K. (1986) *Curr. Genet.* **10**, 463-467.
- Duvell, A., Hessberg-Stutzke, H., Oeser, B., Rogmann-Backwinkel, P., and Tudzynski, P. (1988) *Mol. Gen. Genet.* **214**, 128-134.
- Akins, R.A., Kelley, R.L., and Lambowitz, A.M. (1986) *Cell* **47**, 505-516.
- Akins, R.A., Kelley, R.L., and Lambowitz, A.M. (1989) *Mol. Cell. Biol.* **9**, 678-691.
- Pande, S., Lemire, E.G., and Nargang, F.E. (1989) *Nucleic Acids Res.* **17**, 2923-2942.
- Schulte, U. and Lambowitz, A.M. (1991) *Mol. Cell. Biol.* **11**, 1696-1706.
- Akins, R.A. and Lambowitz, A.M. (1984) *Proc. Natl. Acad. Sci., U.S.A.* **81**, 3791-3795.

27. Kennell, J.C. and Lambowitz, A.M. (1989) *Mol. Cell. Biol.* **9**, 3603–3613.
28. Collins, R.A. and Lambowitz, A.M. (1981) *Curr. Genet.* **4**, 131–133.
29. Lemire, E.G. and Nargang, F.E. (1986) *J. Biol. Chem.* **261**, 5610–5615.
30. Chow, T.Y.-K. and Fraser, M.J. (1983) *J. Biol. Chem.* **258**, 12010–12018.
31. Collins, R.A., and Lambowitz, A.M. (1983) *Plasmid* **9**, 53–70.
32. Taylor, J.W., Smolich, B.D., and May, G. (1986) *Evolution* **40**, 716–739.
33. Kubelik, A.R., Kennell, J.C., Akins, R.A., and Lambowitz, A.M. (1990) *J. Biol. Chem.* **265**, 4515–4526.
34. Robison, M.M., Royer, J.C. and Horgen, P.A. (1991) *Curr. Genet.* **19**, 495–502.
35. Mannella, C.A., Goewert, R.R., and Lambowitz, A.M. (1979) *Cell* **18**, 1197–1207.
36. Almasan, A. and Mishra, N. C. (1990) *Nucleic Acids Res.* **18**, 5871–5872.
37. Gross, S.R., Levine, P.H., Metzger, S., and Glaser, G. (1989) *Genetics* **121**, 693–701.
38. Gross, S.R., Mary, A. and Levine, P.H. (1989) *Genetics* **121**, 685–691.
39. Lambowitz, A.M. (1989) *Cell* **56**, 323–326.