Extraordinary Ribosomal Spacer Length Heterogeneity in a Neotyphodium Endophyte Hybrid: Implications for Concerted Evolution

Austen R. D. Ganley and Barry Scott

Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand

Manuscript received April 13, 1998
Accepted for publication August 22, 1998

ABSTRACT

An extraordinary level of length heterogeneity was found in the ribosomal DNA (rDNA) of an asexual hybrid Neotyphodium grass endophyte, isolate Lp1. This hybrid Neotyphodium endophyte is an interspecific hybrid between two grass endophytes, *Neotyphodium lolii*, and a sexual form, *Epichlöe typhina*, and the length heterogeneity was not found in either of these progenitor species. The length heterogeneity in the hybrid is localized to the intergenic spacer (IGS) and is the result of copy-number variation of a tandemly repeated subrepeat class within the IGS, the 111-/119-bp subrepeats. Copy number variation of this subrepeat class appears to be a consequence of mitotic unequal crossing over that occurs between these subrepeats. This implies that unequal crossing over plays a role in the concerted evolution of the whole rDNA. Changes in the pattern of IGS length variants occurred in just two rounds of single-spore purification. Analysis of the IGS length heterogeneity revealed features that are unexpected in a simple model of unequal crossing over. Potential refinements of the molecular details of unequal crossing over are presented, and we also discuss evidence for a combination of homogenization mechanisms that drive the concerted evolution of the Lp1 rDNA.

CONCERTED evolution is the term used to describe the unusual evolutionary behavior of multigene families whose genes show a great deal of similarity to each other within an array and within a species but accumulate differences between species. This was first demonstrated in the ribosomal multigene family (the rDNA) in Xenopus by Brown *et al.* (1972). This ability of individual repeats in a multigene family to evolve in concert rather than independently is believed to result from a process that is able to homogenize all the repeats in an array (Dover 1982), and this has been directly demonstrated in lizards (Hillis *et al.* 1991) and cotton (Wendel *et al.* 1995). While the concept of concerted evolution resulting from homogenization of the repeat arrays in multigene families is widely accepted, the precise *modus operandi* of the homogenizing mechanism(s) is not well understood (for recent reviews see Elder and Turner 1995; Li 1997).

The two mechanisms most commonly invoked as responsible for homogenization are gene conversion (Edelman and Gally 1970; Nagylaki and Petes 1982) and unequal crossing over (Smith 1973; Ohta 1976). These mechanisms, which fall under the general term "recombination," are proposed to drive a single repeat unit within an array to fixation, with selection presumably removing unfit genes that spread through the array. Gene conversion involves the unidirectional transfer of information from one DNA duplex to another, while unequal crossing over results in reciprocal exchange of information between two DNA duplexes. Much of our understanding of these processes comes from work done in fungi.

The relative roles of unequal crossing over and gene conversion in homogenization are uncertain, and resolution of this debate has been hampered by difficulties in distinguishing these mechanisms experimentally with such a large number of essentially identical genes. Also, unequal crossing over and gene conversion are believed to be mechanistically linked (Holliday 1964; Meselson and Radding 1975; Szostak *et al.* 1983), with gene conversion often being associated with crossing over of flanking markers. However, the isolation of mutants that affect either gene conversion or unequal crossing over demonstrates that they are under some degree of independent control (reviewed in Whitehouse 1982). Recombination can be meiotic or mitotic, and these types also appear to be under at least partially independent control (reviewed in Orr-Weaver and Szostak 1985). Recombination events can occur between homologous chromosomes (allelic or classical recombination), between repeats on nonhomologous chromosomes (heterochromosomal recombination), between sister chromatids (sister chromatid recombination), and even between repeats on the same chromatid (intrachromatid recombination); the latter two processes together are known as intrachromosomal recombination (from Jinks-Robertson and Petes 1993). Intrachromosomal recombination is believed to be the most

Corresponding author: Barry Scott, Institute of Molecular BioSciences, Massey University, Private Bag 11222, Palmerston North, New Zealand. E-mail: d.b.scott@massey.ac.nz

important from a concerted evolution perspective, occurring at a higher rate than the other recombination events (*e.g.*, Petes and Botstein 1977; Petes 1980; Schlötterer and Tautz 1994).

We have been investigating the structure and composition of the rDNA in a group of Neotyphodium grass endophytes from the family *Clavicipitacea*, which includes the fungus responsible for St. Anthony's Fire (ergotism) from contaminated rye. Neotyphodium endophytes are asexual filamentous ascomycetes that form mutualistic symbiotic relationships with pasture grasses, producing a range of secondary alkaloids, including a tremorgenic mycotoxin responsible for the neurotoxic disorder in grazing mammals, ryegrass staggers. Molecular phylogenetic studies indicate that the clavicipitaceous endophytes evolved from the teleomorphic (sexual) choke grass pathogen, Epichlöe (Schardl *et al.* 1991). These sexual forms exhibit pathogenic symptoms through the production of external stroma on the inflorescences of their hosts. Stromal production, which represents the sexual stage of the fungus, prevents maturation of the host inflorescences and leads to sterility, a phenomenon known as "choking." This is in contrast to the Neotyphodium species, which are asexual and entirely endophytic and are disseminated via the host grass seed.

It has been shown recently that several of these Neotyphodium endophytes are interspecific hybrids. Several independent hybridization events appear to have occurred between various sexual Epichlöe species and asexual Neotyphodium species, presumably through hyphal fusion followed by nuclear fusion after dual infection of one plant (Tsai *et al.* 1994). The endophyte we used in this study arose through an interspecific hybridization event between a sexual taxon (*Epichlöe typhina*) and an asexual taxon (*Neotyphodium lolii-Lolium perenne* taxonomic grouping 1 or LpTG-1) that associate with perennial ryegrass (*L. perenne*). The resultant hybrid has been designated taxonomic grouping LpTG-2 (*E. typhina* × *N. lolii*; Schardl *et al.* 1994). Lp1, an isolate from the LpTG-2 hybrid group, and its two progenitors, *E. typhina* isolate E8 and *N. lolii* isolate Lp5 (the extant isolate most closely resembling the true progenitor), were used to investigate the mechanisms of homogenization in the concerted evolution of the rDNA.

The rDNA in most eukaryotes (including fungi) consists of a series of repeating units containing the 18S, 5.8S, and 28S rRNA (*rrn*) genes (Long and Dawid 1980). These units are arranged in a head-to-tail, tandem array with internal transcribed spacers (ITS) separating the *rrn* genes within a unit and an intergenic spacer (IGS) separating adjoining units. The 5S *rrn* gene repeats are usually located separately (Long and Dawid 1980). The *rrn* genes show a remarkable amount of conservation across a wide range of species, while the spacer elements diverge more rapidly. The difference in evolutionary rates arises because the spacers can essentially "hitchhike" with the more functionally constrained *rrn* genes during the homogenization process (Smith 1973; Kellogg and Appels 1995). The IGS is normally maintained at a constant length in the rDNA within a species, but in some species, it fluctuates in length between populations, individuals, and even arrays. In many cases where the structure of the IGS has been determined, it has been found to contain small, tandem subrepeats, and in a number of instances, the length variation of the IGS results from variation in the number of these IGS subrepeats (see discussion). Here we demonstrate unequal crossing over that occurs within the IGS of the rDNA in the Neotyphodium grass endophyte hybrid Lp1. We show that significant changes in repeat variants can arise in just a few generations, and that completely different patterns of length variants arise in a few years.

MATERIALS AND METHODS

Strains and growth conditions: Fungal isolates, λ clones, and plasmids used in this study are listed in Table 1. Fungal isolates were cultured on 2.4% w/v potato dextrose (PD; Difco, Detroit, MI) agar plates at 25°.

Molecular biology techniques: All subcloning was done using the plasmid vector pUC118 (Vieira and Messing 1987), and transformation was performed using *Escherichia coli* strain XL1-Blue (Bullock *et al.* 1987). Standard techniques (restriction digestion, ligation, electrophoretic separation, DNA gel purification, etc.) were done according to the methods of Ausubel *et al.* (1987–1993), unless otherwise described. Purification of PCR products using the PCR Clean Up Kit (Boehringer Mannheim, Mannheim, Germany) and *Exo*III deletion using the Erase-A-Base system (Promega, Madison, WI) were done according to the manufacturers' instructions. DNA sequencing of both plasmid DNA and PCR products using the AmpliCycle (Perkin Elmer, Branchburg, NJ) cycle sequencing kit was performed according to the manufacturer's instructions, except that an annealing temperature of 50° was used. Reactions were carried out in a model FTS-960 thermocycler (Corbett Research, Sydney, Australia).

Library screening and physical mapping: A λEMBL3A genomic library of Lp1 (Collett *et al.* 1995) was screened by plaque hybridization using [α-^{32}P]dCTP-labeled YIp10.4 (Toda *et al.* 1984) by standard procedures (Sambrook *et al.* 1989). Physical mapping was performed using the λ mapping kit (Amersham, Buckinghamshire, England). A DNA fragment containing the 25S rRNA gene (2.4-kb *Hin*dIII/*Bam*HI) was gel purified from YIp10.4. This was radiolabeled with [α-^{32}P]dCTP using the random primer extension method (Feinberg and Vogelstein 1983) for use as a hybridization probe. Primer combinations of ns7-ns8 and its5-its2 (White *et al.* 1990) were used with total DNA to generate PCR products of 350 and 300 bp, respectively. These products were purified using the Magic PCR Preps system (Promega) and were radiolabeled as described above for use as hybridization probes. These were then used as hybridization probes to restriction digests of λPN1 to assign the positions of the *rrn* genes and spacers to the physical map.

DNA extraction: Isolates Lp1 and Lp5 were grown for 6 days and isolate E8 was grown for 4 days in 30 ml of PD liquid broth in flasks on a shaker at 250 rpm at 25°. Mycelia were harvested by filtration through 11-cm Whatman one-filter paper under vacuum, frozen in liquid N_2, and then freeze-dried.

TABLE 1

Fungal isolates, λ clones, and plasmids used in this study

Fungal isolate, λ clone, or plasmid	Relevant characteristics	Source or reference
Fungal isolates		
Neotyphodium sp. (= LpTG-2) Lp1	Neotyphodium sp. isolated from *Lp*[a]	Christensen *et al.* (1993)
Neotyphodium lolii (= LpTG-1) Lp5	*N. lolii* isolated from *Lp*	Christensen *et al.* (1993)
Epichloë typhina (= MP-1) E8	*E. typhina* isolated from *Lp*	Schardl *et al.* (1991)
Phage λ clones		
λEMBL3A	λ(*Aam 32 Bam1*) *sbh1*λ*1*° *b189* (polycloning site *int29 ninL44trpE* polycloning site) KH54 *chiC srl*λ*4*° *nin5 srl*λ*5*°	Frischauf *et al.* (1983)
λPN1	λEMBL3A clone containing rDNA from Lp1	This study
Plasmids		
YIp10.4	*Schizosaccharomyces pombe* rDNA unit clone	Toda *et al.* (1984)
pPN49	pUC118 containing 5.6-kb *Sal*I *rrn* coding region fragment from λPN1	This study
pPN50	pUC118 containing 4.1-kb *Sal*I IGS fragment from λPN1	This study

[a] *Lp, Lolium perenne.*

Total fungal DNA was prepared from the lyophilized mycelia as described previously (Brownlee 1988).

Southern blotting: Southern transfers were carried out according to Ausubel *et al.* (1987–1993). Probes were radiolabeled to a high specific activity with [α-^{32}P]dCTP using the high prime random priming kit (Boehringer Mannheim), and the unincorporated nucleotides were removed with a Sephadex G-50 column (ProbeQuant G-50 micro column; Pharmacia Biotech, Piscataway, NJ). DNA hybridizations were carried out in 10× Denhardt's solution (3× SSC) at 65° for 16 hr. Three sets of washes were performed, each at room temperature in 2× SSC for 15 min. Detection and stripping of the Southern blots were performed as described in Ausubel *et al.* (1987–1993).

Single-spore purification: Single-conidiospore-purified cultures of Lp1 were generated as follows. Lp1 was cultured onto a fresh PD plate and grown for 2 weeks at 25°. Conidiospore solutions were prepared by immersing a mycelial block from the plate culture in sterile H_2O. This solution was plated onto PD agar 2% (w/v) plates and allowed to germinate for 48 hr at 25°. Germinating conidiospores were identified by microscopy and then picked and patched onto fresh PD plates.

PCR analysis: All PCR reactions were carried out in a final volume of 25 μl containing 10 mm Tris-HCl, 1.5 mm $MgCl_2$, 50 mm KCl, 50 μm of each dNTP, 200 nm of each primer, 1 unit of *Taq* DNA polymerase (Boehringer Mannheim), and 10 ng of genomic DNA. The temperature regime used was as follows: 2 min at 94°; 25 cycles of 30 sec at 94°, 30 sec at the temperature indicated, and 1 min at 72°; and 5 min at 72°. For the primer combinations nts1 (5′-CGGCTCTTCCTATCA TACCGAAG-3′) with nts2 (5′-GACTCCCCTCGGGATTAGCA TAG-3′) and nts7 (5′-TGCGGGTGCGCTATCGAGATG-3′) with nts8 (5′-GCAAATCACAGTCACCAGCGG-3′), a 57° annealing temperature was used. For the primer combination nts3 (5′-TCTTGCAGACGTCTACTCCGTG-3′) with nts4 (5′-GAG ACAAGCATATGACTAC-3′), a 55° annealing temperature was used, and DMSO was added to a final concentration of 2% (v/v). Reactions were carried out in a model FTS-960 thermocycler (Corbett Research).

Multivariant repeat PCR (MVR-PCR) analysis: For the mapping of the 111-/119-bp subrepeats, 18-mer oligonucleotides specific to each of the subrepeat types that covered the variable region of the subrepeats at the 3′ end of the primer were designed in both directions. In addition, an arbitrary 17-mer tail was constructed at the 5′ end of each primer to produce the final 35-mer repeat-specific primers, and a primer consisting of this sequence alone, the TAG primer (5′-TTTG TCCGCTCGGTTGC-3′), was also constructed. The repeat-specific primer sequences are as follows: 111-L is 5′-TTTGTCCG CTCGGTTGCCGCGGGCAGAGTGGTGCC-3′, 111-R is 5′-TT TGTCCGCTCGGTTGCGCCCATCCCACCACTCTG-3′, 119-L is 5′-TTTGTCCGCTCGGTTGCTCAGAGTGGTGTCCTCGG-3′, and 119-R is 5′-TTTGTCCGCTCGGTTGCCGCCCATCC AACCCGAGG-3′. The anchor primers, located to the left and right of the 111-/119-bp subrepeat array, are previously designed primers, where anchor-L is nts7, and anchor-R is nts4 for genomic DNA and is the pUC118 forward primer for pPN50 DNA.

All MVR-PCR reactions consisted of one initial cycle that was carried out in a final volume of 50 μl containing 10 mm Tris-HCl, 1.5 mm $MgCl_2$, 50 mm KCl, 5% (v/v) DMSO, 50 μm of each dNTP, 10 nm of the repeat specific primer, 300 nm of the appropriate anchor primer, 3 units of High Fidelity *Taq* DNA polymerase (Boehringer Mannheim), and 10 ng of genomic DNA or 0.2 ng of pPN50 DNA. The temperature regime was 3 min at 92°, 30 sec at 60°, and 5 min at 70°. After this, 300 nm TAG primer and an additional unit of High Fidelity *Taq* DNA polymerase were added to the reaction, and the reactions were put back into the thermocycler with the following temperature regime: 2 min at 92°; 25 cycles of 30 sec at 92°, 30 sec at 60°, and 3 min at 70°; and 5 min at 70°. All reactions were carried out in a model FTS-960 thermocycler (Corbett Research). Reactions were then fractionated on agarose gels.

RESULTS

Cloning of an Lp1 rDNA unit: A λEMBL3A genomic library of Lp1 (Collett *et al.* 1995) was screened for rDNA clones with a 10.4-kb *Hin*dIII rDNA probe from *Schizosaccharomyces pombe* (YIp10.4; Toda *et al.* 1984), and one clone, λPN1, was selected for further analysis. A physical map of this clone was created using the restriction enzymes *Sal*I and *Eco*RI (Figure 1). A 2.4-kb *Hin*dIII/*Bam*HI fragment from YIp10.4 encoding the *S. pombe* 25S *rrn* gene was used as a probe to Southern blots of λPN1 DNA to assign the position of this gene

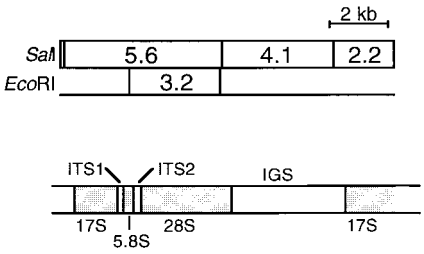

Figure 1.—Map of the ribosomal clone from Lp1, λPN1. Fragments produced by *Sal*I and *Eco*RI digestion are indicated. A diagrammatic representation of the organization of the *rrn* spacers (above) and coding regions (below) in λPN1 is shown below the physical map. Shaded boxes indicate *rrn* coding regions. A complete rDNA unit is defined by two *Sal*I sites, one cutting just before the 5′ end of the 18S *rrn* gene, and the other cutting just after the 3′ end of the 28S *rrn* gene, producing 4.1- and 5.6-kb fragments. The next *Sal*I site begins the next rDNA unit in the array. The 2.2-kb *Sal*I fragment is a partial copy of the 5.6-kb *Sal*I fragment interrupted by the *Sal*I site from the λEMBL3A multicloning site. The other λEMBL3A *Sal*I site is the left-hand-most site in the map. There are also two *Eco*RI sites in the Lp1 rDNA. One cuts within the 5.8S *rrn* gene, and the other at the 3′ end of the 28S *rrn* gene, producing the 3.2-kb fragment indicated. The rDNA unit is truncated before the next *Eco*RI site, which would be expected to produce a 6.5-kb fragment. The 5.6- and 4.1-kb *Sal*I fragments were cloned into pUC118 to give pPN49 and pPN50, respectively.

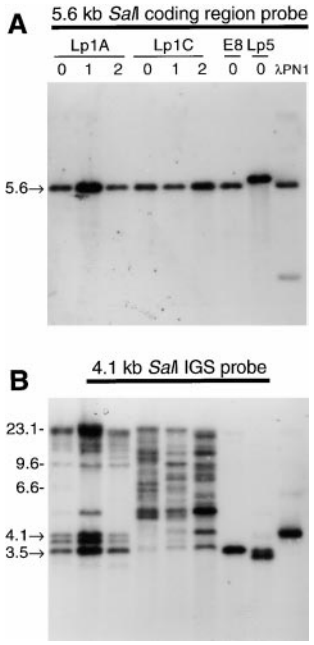

Figure 2.—Length heterogeneity in the rDNA of Lp1. Genomic DNA was digested with *Sal*I, fractionated through a 0.7% agarose gel, and transferred to a nylon membrane. The resulting Southern blot was probed with the two *Sal*I fragments from the Lp1 rDNA unit clone, λPN1. (A) The 5.6-kb coding region probe reveals a single hybridizing band. This is the size of the probe, as indicated, except in the case of Lp5. (B) The 4.1-kb IGS probe reveals a remarkable amount of length variation in the rDNA despite being linked to the DNA in A *in vivo*. The smallest IGS length (3.5 kb) and the IGS length equivalent to the clone (4.1 kb) are indicated by arrows. The order of the lanes in B is identical to that in A. *Hin*dIII size markers from λ DNA are also shown. The original isolate and first and second round single-spore isolates are indicated by 0, 1, and 2, respectively.

to the map. PCR products generated using "universal" fungal primers to the 18S *rrn* gene (ns7 and ns8) and the 18S-5.8S ITS-1 (its5 and its2; White *et al.* 1990) were used as probes to assign the positions of these regions to the map. Positions of the 5.8S *rrn* gene and the ITS-2 region have been inferred from DNA sequencing of these regions (Schardl *et al.* 1994). A map of the Lp1 rDNA repeat unit was constructed from these results (Figure 1). The 5.6- and 4.1-kb *Sal*I fragments from λPN1 were subcloned into pUC118 to generate pPN49 and pPN50, respectively.

Length heterogeneity discovered with a ribosomal probe: To ensure that λPN1 was representative of the genomic rDNA organization, Lp1 genomic DNA was cleaved with *Sal*I, and a Southern blot was probed with the inserts from pPN49 and pPN50. The insert from pPN49, containing the three *rrn* genes on a 5.6-kb *Sal*I fragment (the 5.6-kb coding region probe), hybridized to a 5.6-kb band as expected (Figure 2A). When the insert from pPN50, containing the IGS region on a 4.1-kb *Sal*I fragment (the 4.1-kb IGS probe), was used to probe the same Southern blot, it hybridized to a multitude of bands ranging in size from 3.5 to >20 kb (Figure 2B), including a band the size of the subcloned fragment. All four different cultures of the Lp1 isolate maintained in our laboratory (Lp1A, Lp1C, Lp1D, and Lp1F) exhibited distinct banding patterns when probed with the 4.1-kb IGS probe (results shown for Lp1A and Lp1C in Figure 2B).

This length variation in the Lp1 rDNA is not observed in the progenitor isolates Lp5 and E8 (Figure 2B), although additional bands are observed in these two isolates. In Lp5, three bands, ranging from 3.5 to 3.7 kb, hybridize to the 4.1-kb IGS probe. This limited length heterogeneity is qualitatively different than that seen with Lp1. In the case of E8, there is a strongly hybridizing band of ~3.7 kb and a weaker band just above this.

Aside from some faint hybridization of ~3.2 kb in Lp1, no bands smaller than the two progenitors are seen. Therefore, the hybridizing bands seen in Lp1 are at least as great in size as the bands present in the two progenitors.

Heterogeneity occurs within the rDNA cluster: To ascertain whether the variation in rDNA length is the result of intercellular differences between rDNA clusters or occurs within an rDNA cluster, different laboratory cultures were taken through two rounds of single-conidiospore purification, with genomic DNA extracted after each round. Asexual isolates, including the interspecific hybrids, retain the ability to produce conidiospores. These spores are uninucleate (Schardl *et al.* 1994), allowing pure cultures of a homogenous nuclear composi-

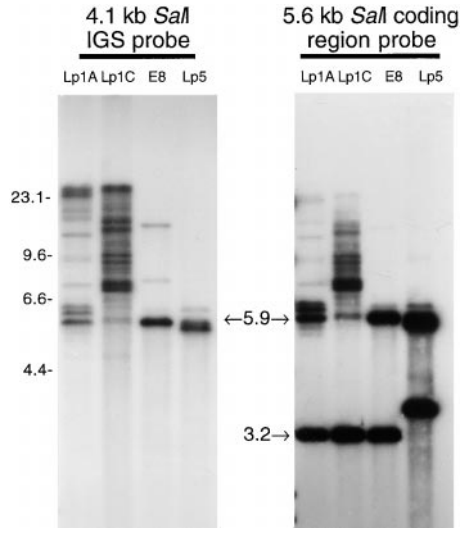

Figure 3.—Length heterogeneity in the rDNA is localized to the spacer. Genomic DNA was digested with *Eco*RI and fractionated through a 0.7% agarose gel and transferred to a nylon membrane. The resulting Southern blot was probed with the two *Sal*I fragments used in Figure 2. The length heterogeneity observed in Figure 2B is also evident in this blot. The pattern of bands is the same as for the *Sal*I genomic digests, but the sizes have increased by 2.5 kb, corresponding to the extra 18S-containing *Eco*RI/*Sal*I fragment. The 5.6-kb coding region probe gives a pattern of hybridizing bands identical to that seen with the 4.1-kb IGS probe, excluding some of the very high molecular weight bands. In addition, a 3.2-kb band is seen, corresponding to the 28S-containing *Eco*RI fragment. The smallest IGS length (5.9 kb), equivalent to the 3.5-kb band in Figure 2B, is indicated by an arrow. *Hin*dIII size markers from λ DNA are also shown. Lp1A and Lp1C genomic DNA is from first-round single-spore-purified cultures.

tion to be generated from mycelia by the germination of a single spore. The DNA from these laboratory cultures was digested with *Sal*I and a Southern blot probed with the 4.1-kb IGS and 5.6-kb coding region probes (results shown for Lp1A and Lp1C in Figure 2). Each laboratory culture retained its unique banding pattern through the single-sporing rounds with the 4.1-kb IGS probe. Two conclusions can be drawn. First, the rDNA length variants are present within the rDNA cluster rather than resulting from intercellular rDNA differences. Second, some changes in the banding patterns were observed through the single sporing, indicating that the mechanism(s) causing the length heterogeneity is still active (*e.g.*, Figure 2B, Lp1A lanes marked 1 and 2).

Heterogeneity localized to the intergenic spacer: The banding pattern seen in Figure 2B may result from heterologous hybridization of the 4.1-kb IGS probe. Therefore, Lp1 genomic DNA was cleaved with *Eco*RI, and a Southern blot was probed with the 4.1-kb IGS and 5.6-kb coding region probes (Figure 3). The physical map of λPN1 (Figure 1) shows that an *Eco*RI genomic digest should produce two rDNA fragments (excluding the flanking fragments), one corresponding approximately to the 5.8S and 28S genes (3.2 kb in size) and the other corresponding to the 18S gene and the IGS (6.5 kb in size). Probing with the 4.1-kb IGS probe produced the same general banding pattern seen in the *Sal*I digestion, with the bands all greater in size by 2.5 kb (the size of the 18S gene) than the bands seen in Figure 2B. The 5.6-kb coding region probe hybridized to a 3.2-kb band as expected, and it also hybridized to the same multitude of bands that the 4.1-kb IGS probe hybridized to. The heterogeneous banding pattern is the result of linkage of the 18S gene and the IGS in the *Eco*RI digest, with the 5.6-kb coding region probe hybridizing to the 18S moiety. This demonstrates that the heterogeneous banding pattern is not the result of heterologous hybridization of the 4.1-kb *Sal*I IGS probe, but that it is length variation of the IGS. It also confirms that the *Sal*I bands seen in Figure 2, A and B, are linked *in vivo*, and that the multitude of bands seen in Figure 2B are ribosomal in origin. The same pattern of hybridization is seen with the two progenitors, E8 and Lp5, but without the multitude of hybridizing bands, as expected. These results also rule out the possibility that the hybridization pattern is the result of an unlikely experimental artefact, such as "star" activity of the restriction enzymes or methylation of the rDNA.

Chromosomal karyotype analysis of the different laboratory cultures and their respective single-spore-purified isolates using pulsed-field gel electrophoresis revealed no differences in their chromosomal banding patterns (A. R. D. Ganley, unpublished results). Southern blots of restriction enzyme digests probed with an Lp1 *pyr*4 clone (pMC11; Collett *et al.* 1995) and an Lp1 *hmg* clone (Dobson 1997) did not reveal any length heterogeneity within any of the laboratory cultures. Thus, the length heterogeneity observed with the ribosomal probes does not appear to be a general feature of the Lp1 genome, nor is it a result of gross chromosomal rearrangements.

IGS contains subrepeat elements: To investigate the nature of the length heterogeneity, the 4.1-kb IGS insert from pPN50 was sequenced. A set of nested deletions of pPN50 was created using exonuclease III, and these were cloned and sequenced. Primers were designed from the sequence obtained to fill gaps in the sequence. This resulted in a complete single-stranded sequence (barring some repetitive elements; see below for details) for the 4.1-kb IGS clone. Sequence was also obtained for the edges of the 5.6-kb coding region insert from pPN49.

Analysis of the IGS sequence revealed two subrepeat classes (Figure 4). The first, termed the 40-bp subrepeat class, is a relatively heterogeneous class, with a core consensus of 40 bp (Figure 4). The individual repeats of this class are organized in a head-to-tail tandem array, with eight repeats present in the 4.1-kb IGS clone. This

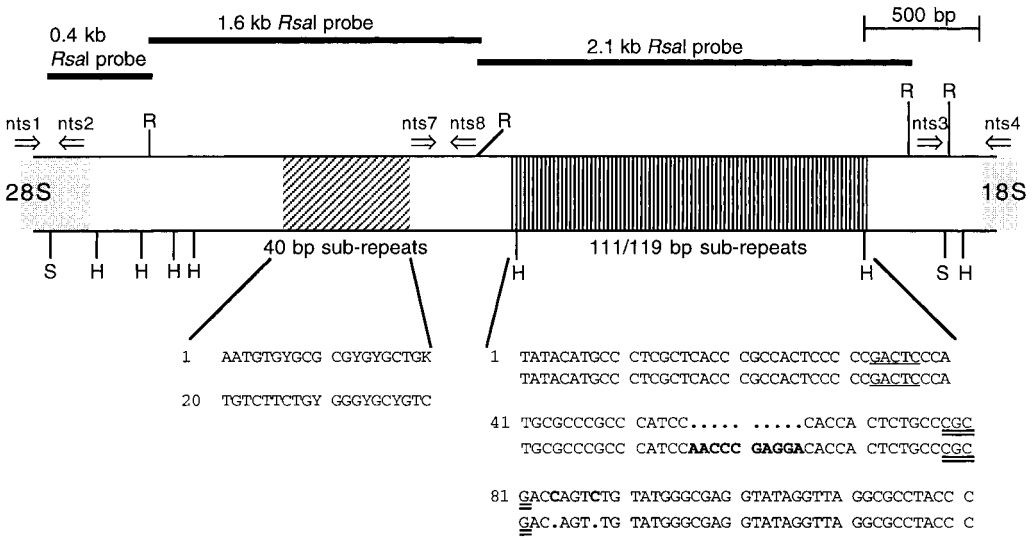

Figure 4.—Organization of the IGS in Lp1 rDNA. The positions and relative sizes of the 40- and 111-/119-bp subrepeat arrays are shown. Below are the consensus sequences for the subrepeats. The ambiguity characters in the 40-bp subrepeat consensus sequence, Y (C or T) and K (G or T), indicate that these nucleotides are present in that position in roughly equal proportions. The sequence in bold indicates differences between the 111- and 119-bp subrepeat consensus sequences. The *Hin*fI site is underlined, and the *Tha*I site is double underlined. The primer pairs nts1-nts2, nts3-nts4, and nts7-nts8 used for sequence comparisons are indicated above the IGS diagram. Restriction sites are as follows: H, *Hin*fI; R, *Rsa*I; and S, *Sal*I. Only the terminal *Hin*fI sites in the 111-/119-bp subrepeats are shown. The three major IGS fragments defined by *Rsa*I (0.4, 1.6, and 2.1 kb) that were used as probes in Figure 5 are shown as thick lines above the figure. The sequences for the PCR products and the 111-/119-bp subrepeats have been deposited in GenBank under accession numbers AF049246 and AF049673–AF049681.

subrepeat class is characterized by alternating pyrimidine-purine residues.

The second subrepeat class is termed the 111-/119-bp subrepeat class and is composed of two very closely related subrepeats, one 111 bp in length (GenBank accession number AF049675) and the other 119 bp in length (GenBank accession number AF049676). These show a high level of identity to each other and to themselves (Figure 4). They are also organized in a head-to-tail tandem array. The subrepeats of this class are GC rich, containing on average 65% GC. The 4.1-kb IGS clone was shown to contain 14 repeats (see below).

The junction between the 3′ end of the 28S gene (in pPN49) and the 5′ end of the IGS (in pPN50) was spanned using primers designed from the sequence obtained (nts1 and nts2). PCR amplification of Lp1 genomic DNA using this primer combination produced a product of the expected size, 210 bp, and sequencing confirmed that this product contained a *Sal*I site at the appropriate location (GenBank accession number AF049679). This is further confirmation that the 4.1-kb IGS and 5.6-kb coding region fragments are linked *in vivo*. The nts1-nts2 PCR product is wholly 28S *rrn* sequence, although we have not precisely determined the 3′ end of the 28S gene. The junction between the 3′ end of the IGS (in pPN50) and the 5′ end of the 18S gene (in pPN49) was spanned using a primer that is the reverse complement of ns1 (White *et al.* 1990; referred to as nts4) and a primer designed from the sequence of the 4.1-kb IGS clone (nts3). PCR amplification using this primer combination with Lp1 genomic DNA produced a product of the expected size, 479 bp, and sequencing confirmed that this product contained a *Sal*I site at the appropriate location (GenBank accession number AF049246).

Digestion of the 111-/119-bp subrepeats abolishes heterogeneity: Length variation in the IGS of other organisms results from variation in the number of subrepeats, and we suspected that the 111-/119-bp subrepeats were responsible for the length heterogeneity in Lp1. We identified two restriction enzymes (*Hin*fI and *Tha*I) that cleave these subrepeats once (Figure 4). If copy-number variation of these subrepeats is responsible for the length heterogeneity, cleaving them with either *Hin*fI or *Tha*I should abolish the heterogeneity, leaving a high-copy-number band the size of the subrepeats.

To simplify the analysis of the results, we identified a restriction enzyme (*Rsa*I) that cleaves the IGS into three smaller fragments suitable for probes (Figure 4). Genomic DNA was cleaved with *Hin*fI and *Tha*I, and the Southern blots were probed with the three *Rsa*I subfragments derived from the 4.1-kb IGS fragment. The results for *Hin*fI are shown in Figure 5. None of the three *Rsa*I probes reveals any evidence of the length heterogeneity seen in Figure 2B. The bands present (except the 270-bp band, see below) are all of the sizes predicted from the sequence of the 4.1-kb IGS clone. The probe covering the region that includes the 111-/119-bp subrepeats is the 2.1-kb *Rsa*I probe. This shows no evidence of length heterogeneity after *Hin*fI digestion for any Lp1 culture studied, and, as predicted, there is a strongly hybridizing band the size of a single subrepeat (~115 bp). In one of the Lp1 laboratory cultures (Lp1A), an unexpected band is found (1.1 kb; marked with an asterisk). This appears to be a length polymorphism in the spacer that is not the result of the 111-/119-bp subrepeats. However, it is not able to explain the level of the IGS length heterogeneity seen in Figure 2B. The precise nature of this polymorphism is not clear. Digestion with *Tha*I also abolished all evi-

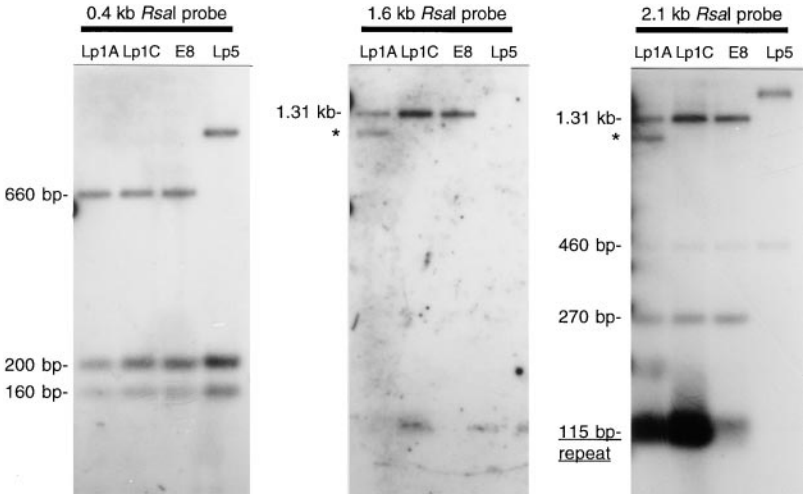

Figure 5.—Digestion of the 111-/119-bp IGS subrepeats abolishes the length heterogeneity. Genomic DNA was digested with *Hin*fI and fractionated through a 3.0% agarose gel and transferred to a nylon membrane. The resulting Southern blot was probed with the three *Rsa*I IGS subfragments shown in Figure 4. All three probes produce the expected pattern of hybridizing bands, and their sizes are indicated (refer to Figure 4). The two exceptions are the 270-bp band seen with the 2.1-kb *Rsa*I probe and the band marked with as asterisk (see text). Several bands smaller than 200 bp are not visible in all these exposures. A strongly hybridizing band at ~115 bp is observed with the 2.1-kb *Rsa*I subfragment. This band corresponds to the 111-/119-bp subrepeats, and the strong hybridization indicates high copy number. None of these *Rsa*I probes show any evidence of length heterogeneity. Lp1A and Lp1C genomic DNA is from first-round single-spore-purified cultures.

dence of length heterogeneity (A. R. D. Ganley, unpublished results). Abolition of length heterogeneity with enzymes that cut the 111-/119-bp subrepeats demonstrates that the IGS length heterogeneity is the result of copy-number variation of the 111-/119-bp IGS subrepeats.

Probing *E. typhina* isolate E8 genomic DNA with the 2.1-kb *Rsa*I fragment produces the same bands seen in Lp1, except that the intensity of hybridization of the band at ~115 bp is not as strong (Figure 5). The *N. lolii* isolate Lp5 produces a somewhat different pattern of bands, and there does not appear to be any hybridization at ~115 bp (Figure 5). The IGS structure in Lp5 appears to differ from that found in Lp1 and E8 (see below).

Arrangement of the 111-/119-bp subrepeats in the IGS: Information on the number and distribution of repeats within an array can give insights into the processes that are shaping the array. We used MVR-PCR (Jeffreys *et al.* 1991) to determine the order of the two types of subrepeats in the 111-/119-bp subrepeat array (Dover *et al.* 1993). We found that the High Fidelity *Taq* DNA polymerase (Boehringer-Mannheim) used in the PCR reactions had difficulty in traversing these subrepeats. Hence, two MVR-PCR reactions, initiating from each end of the subrepeat array, were needed to determine the order of the entire array in the clone. The strategy we used to determine the order of the subrepeat array is shown schematically in Figure 6, and we carried this out for both the Lp1 clone (pPN50) and genomic DNA. The order of subrepeats for each half of the array is determined by reading the order of bands up the pairs of lanes in the gels shown in the lower part of Figure 6 for each L-R subrepeat-specific primer set. Combining the results from each end of the array and finding the overlap gives the complete order of subrepeats in the clone (shown in Figure 6). No obvious pattern of organization of the two subrepeats is evident.

The 111-L subrepeat-specific primer is not specific for the 111-bp subrepeats, but it amplifies both the 111- and 119-bp subrepeats equally well. This is likely to be a consequence of the primer sequence, as the only bases in this primer that are specific for the 111-bp subrepeat are the two 3′-most bases. This difference does not appear to be sufficient to distinguish the two subrepeats under the PCR conditions used. Raising the annealing temperature abolished amplification (data not shown), presumably because of the left anchor primer failing to anneal. This lack of specificity does not prevent the ordering of the subrepeats because the specificity of the 119-L subrepeat-specific primer clearly shows the order.

The results for the MVR-PCR with genomic DNA are also shown in the gels in Figure 6. The results for genomic DNA give some idea about the conservation or lack thereof of subrepeat order in the IGS within the rDNA cluster. The bands from the right-hand side of the subrepeat array can be ordered for about six subrepeats, and this order is the same as in the clone. The specificity of banding is not as clear as in the clone, indicating that some heterogeneity of subrepeat order exists among the population of IGS, but, nevertheless, the order can be determined. Conversely, the results for the left-hand side of the subrepeat array do not resemble the clone, and there does not appear to be any clear ordering of the subrepeats at this edge of the array. Once again, the lack of specificity of the 111-L subrepeat-specific primer is not likely to confound the results, as gaps in the 119-L subrepeat-specific primer ladder would be expected. Instead, the 119-L subrepeat-specific primer anneals to many more subrepeats with genomic DNA as the template than with the clone. This indicates that the population of IGS has considerable variation in the order of subrepeats at this end of the array. Another feature not found with the clone is the presence of an extra band in the ladder of subrepeats on the left-hand side of the array (marked with an asterisk in Figure 6).

This appears to result from two truncated subrepeats appearing in the array, whose combined size is approximately that of a single, full-length subrepeat. These truncated subrepeats would not contain *Hin*fI sites, and this appears to be the origin of the 270-bp band seen in Figure 5 that is not predicted from the IGS sequence.

Lp1 rDNA is derived from E8: To determine how the sequence of the rDNA in Lp1 related to the hybrid nature of Lp1, PCR was performed with genomic DNA from the two progenitor isolates and with genomic DNA from Lp1. We used the nts1-nts2, nts3-nts4, and nts7-nts8 primer combinations (refer to Figure 4). The resulting PCR products were sequenced and compared. Lp1 and E8 had identical sequences for all three PCR products. Comparing the nts1-nts2 PCR products derived from the 3′ end of the 28S *rrn* gene, between Lp1 and Lp5, we found six substitutions and one indel over 205 bp (94.6% identity). Comparing the nts3-nts4 PCR products derived from the 3′ end of the IGS, between Lp1 and Lp5, we found 30 substitutions and three indels over 482 bp (93.7% identity). No product was amplified using nts7-nts8 in Lp5, indicating that this region is sufficiently different in this endophyte to prevent amplification (GenBank accession numbers AF049246, AF049673, AF049674, and AF049677–AF049681). Therefore, the rDNA in Lp1 appears to be exclusively derived from E8, with no evidence of any Lp5 rDNA sequence being found. Restriction enzyme digests and Southern blotting data are all consistent with this conclusion.

DISCUSSION

The results reported here reveal a high level of length heterogeneity in the ribosomal IGS region of the hybrid endophyte isolate Lp1. We have shown that this heterogeneity is intragenomic and is not the result of length differences between cells. The length heterogeneity arises from copy-number variation of a subrepeat class, the 111-/119-bp subrepeats, which are located within the IGS. This copy number variation is likely to be a consequence of unequal crossing over occurring in the register of the 111-/119-bp subrepeats. Unequal crossing over occurs when tandemly repeated elements misalign and then undergo a reciprocal exchange. As a consequence of this misalignment, the two reciprocal products that are formed each contain a different number of repeats from the original molecules, and the extent of misalignment determines the extent of change of copy numbers in the products. IGS length heterogeneity as a consequence of copy-number variation of subrepeats has been reported in a number of organisms previously, including vertebrates (Wellauer *et al.* 1976; Botchan *et al.* 1977; Krystal and Arnheim 1978; Arnheim and Kuehn 1979), insects (Wellauer and Dawid 1978; Schafer *et al.* 1981; Coen *et al.* 1982; Israelewski and Schmidt 1982; Tautz *et al.* 1987), plants (see Rogers and Bendich 1987), and fungi (Martin 1990). However, the extent of the length heterogeneity, the fact that it is clearly intragenomic, and the sudden appearance of the heterogeneity from organisms that do not display such heterogeneity (the two progenitor isolates) make this system particularly interesting.

The unequal crossing over we have demonstrated within the IGS subrepeats is likely to play a role in the concerted evolution of the whole rDNA if there is also unequal crossing over between whole rDNA units. There is no *a priori* reason to suspect that misalignment of the subrepeats occurs with equal crossing over of the

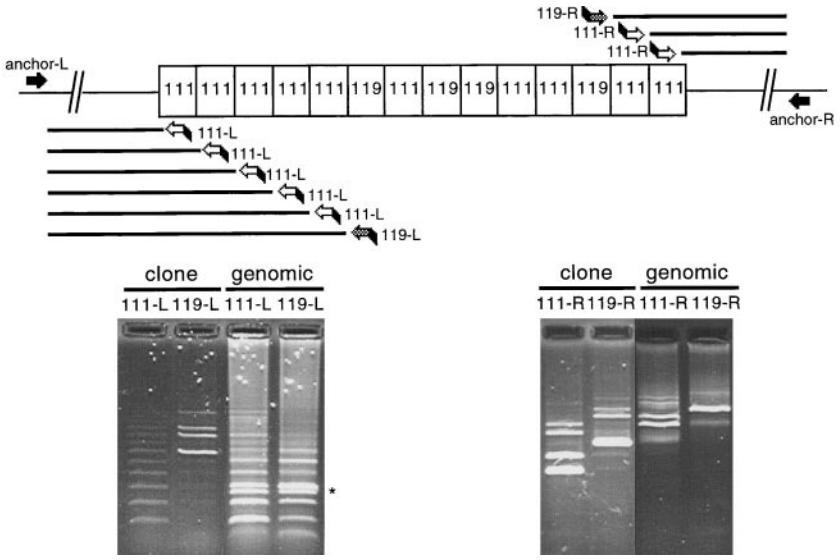

Figure 6.—Multivariant repeat PCR analysis of the 111-/119-bp subrepeats. The upper part of the figure shows the general scheme used for the MVR-PCR analysis. The 111-/119-bp subrepeats are shown in boxes; the orientation is the same as that shown in Figure 4. The first round of PCR amplification is shown schematically from each end of the subrepeat array. The subrepeat-specific primers 111-L, 111-R (open arrows), 119-L, and 119-R (stippled arrows) anneal specifically to the appropriate subrepeat in the array, and they amplify the intervening sequence in combination with the anchor-L or anchor-R primers (solid arrows). The ladder of products from this first round of amplification is shown as thick lines. These products are then stably amplified by further PCR rounds using the TAG primer (shown as a filled tail on the subrepeat-specific primers) and the anchor-L/anchor-R primers. The amplification products from the IGS clone (pPN50) and Lp1 genomic DNA for each half of the subrepeat are shown in the gels in the lower part of the figure. Reading the ladders of subrepeats from the gels gives the order of the subrepeats. The subrepeat order in the upper part of the figure is the order found in the IGS clone. An extra band seen with genomic DNA on the left-hand side of the array is indicated by an asterisk.

rDNA units in the array. Indeed, the size of the rDNA cluster is found to vary in many species, implying that variation in rDNA unit copy number as a result of unequal crossing over is common. Therefore, we propose that the unequal crossing over that generates the copy number variation of the 111-/119-bp subrepeats in the IGS is concomitant with unequal crossing over in the register of the rDNA units and, therefore, plays a homogenizing role in the concerted evolution of the Lp1 rDNA.

These data present an apparent paradox—the process that plays a role in the homogenization of repeats (unequal crossing over) is responsible for generating a high level of heterogeneity in these repeats. This paradox is expected as the homogenization process is working at the sequence level, and the misalignment that drives the sequence homogenization produces heterogeneity at the level of length. So unequal crossing over will tend to spread a particular repeat throughout the array, but, as this repeat spreads, different copies acquire different numbers of subrepeats strictly as a result of the process of spread (Kelly *et al.* 1990).

Although the distinction between mitotic and meiotic recombination is well appreciated, little has been done to assess the relative roles that these different forms of recombination play in concerted evolution. The unequal crossing over we have demonstrated here is strictly mitotic, as Lp1 is an asexual organism (M. Christensen, personal communication). There is evidence, aside from the lack of breakdown of multigene families in asexual organisms, that mitotic recombination plays an important role in concerted evolution. Both unequal crossing over (Szostak and Wu 1980) and gene conversion (Jackson and Fink 1981) occur in mitosis. Interestingly, the rate of mitotic recombination is normally low in the genome, with the rDNA being an exception to this rule (Szostak and Wu 1980). It has even been suggested that meiosis may slow the rate of homogenization that would occur with mitosis alone (Elder and Turner 1995). However, it is likely that the importance of mitosis in concerted evolution is the result of intrachromosomal rather than interchromosomal recombination doing much of the "work" in concerted evolution, as mitotic recombination is primarily intrachromosomal. Our results concur with this idea.

The rate and nature of turnover in the IGS: The rate of turnover caused by unequal crossing over in Lp1 is high. Two rounds of single-spore purification are sufficient to produce noticeable changes in the pattern of IGS lengths (Figure 2B). Furthermore, this turnover is able to produce drastic changes in a relatively short space of time. The two laboratory cultures presented here, Lp1A and Lp1C, were derived from the initial isolate culture Lp1 and were maintained as separate plate cultures for ~4 yr before the DNA used in this study was extracted. In this time, they have evolved their own distinct banding profiles, arising from one initial profile. Other laboratory cultures have evolved their own distinct profiles as well (A. R. D. Ganley, unpublished results). This rapid rate of turnover may explain the remarkable spread of IGS lengths that we observe. The longest IGS lengths must contain at least 200 subrepeat units. This would involve a great number of sequential unequal crossing over events from the original 10–15 subrepeats. The problem is exacerbated if the degree of misalignment allowed in unequal crossing over is small. However, it also remains possible that the long IGS lengths that show up faintly with E8 in Figure 2 may have somehow "seeded" the great number of long IGS lengths found in Lp1.

The changes in the IGS banding pattern through the two rounds of single-spore purification have an unexpected feature. Several of the bands that appear or disappear through the single sporing are strongly hybridizing and, therefore, must represent a number of copies of that particular IGS length. Unequal crossing over between rDNA units will result in the stochastic loss or gain of rDNA units. Therefore, loss of a strongly hybridizing band is likely to represent the loss of a block of rDNA units, all with the same IGS length. This implies that rDNA units with the same IGS lengths are clustered together, and further implies that the strongly hybridizing bands that appear are also clustered. Clustering of length variants may arise when the degree of misalignment in recombination is small. Szostak and Wu (1980) found that the degree of misalignment in yeast rDNA mitotic unequal crossing over was six to eight units, and this corresponds well with that found in *Drosophila melanogaster* 5S RNA unequal crossing over (Samson and Wegnez 1988). Dvorák *et al.* (1987) showed that gene conversion in the IGS subrepeats was distance dependent. These results suggest that clustering may arise as a consequence of the localized mode of action of the homogenization mechanism(s), and Dvorák and Appels (1986), Crease (1995), and Copenhaver and Pikaard (1996) explained clustering along these lines. However, the limited time required by our results to generate such clustered length variants stretches the bounds of plausibility, as several sequential misalignment events would be required in the amount of mitotic growth it takes to generate a fungal disk from a single spore (assuming a length variant arises once and is then amplified). The process also appears to be targeted in the sense that only very few length variants alter their copy number, but those that do then change by a significant number of copies. Reeder *et al.* (1976) found a similar phenomenon in the rDNA of *Xenopus laevis.* They suggested this was a consequence of extrachromosomal amplification of rDNA units and their insertion into the rDNA array. However, this form of amplification, well documented in the oocytes of Xenopus, has not been reported in fungi.

The IGS lengths we see do not represent a clean ladder of bands at ~115-bp intervals, as one might ex-

pect if the length varied as a result of the 111-/119-bp subrepeats. Instead, the number of lengths observed is limited (the cultures presented in Figure 2 contain 8–16 predominant bands). Lp1A in particular shows a very skewed range of IGS lengths, falling almost entirely in either the high- or low-molecular-weight part of the range. Restriction of IGS lengths to a small number of the total possible set is presumably the result of a homogenization process that (stochastically) amplifies this subset of IGS lengths at the expense of others. This homogenization process is unlikely to be unequal crossing over, as this would tend to increase the IGS length variation as long as there was misalignment of the subrepeats. We are then left with the possibility that gene conversion is responsible for the restriction of IGS lengths we observe. Previous workers have proposed a combination of mechanisms to account for homogenization (Williams *et al.* 1989; Linares *et al.* 1994; Crease 1995).

Molecular details of the IGS subrepeat behavior do not conform to the standard model of unequal crossing over: Our lack of understanding of the mechanisms behind homogenization extends to a lack of knowledge of the biochemistry and genetics of these mechanisms. Therefore, systems that provide data on the particular mechanisms of homogenization, such as the one we have studied, may also provide information on the molecular details of these mechanisms.

In repeat arrays shaped by the forces of unequal crossing over, variants that arise and become eliminated from the array do so by being moved to the edges of the array, a phenomenon known as terminal exclusion (Dover *et al.* 1993). This occurs because the crossover point is assumed to be random. The MVR-PCR analysis we performed gives the order of the 111-/119-bp subrepeats in the clone, and the distribution of the presumed "variant" (119-bp) subrepeat toward the center of the array shows that terminal exclusion is not occurring for this IGS. The results in genomic DNA corroborate this. We also find little evidence from the sequence for degraded subrepeats at the edges of the array. Two alternative but not mutually exclusive explanations can be made for the lack of terminal exclusion. Either the 111- and 119-bp subrepeats are both required for some role in the IGS, particularly its concerted evolution, or other forces alongside unequal crossing over are also shaping this subrepeat array.

The incongruity in subrepeat patterns between the right and left sides of the array in genomic DNA (Figure 6) is unexpected. From the assumption of crossing over occurring at a random point, it follows that both ends of the array should behave the same, but they appear not to, as one end resembles the clone and the other does not. Two explanations are possible: either the crossover point is not random but is specifically initiated from one side of the array, or other forces alongside unequal crossing over are involved in the subrepeat array. Although we have presented some circumstantial evidence for mechanisms of homogenization, most probably gene conversion, that occur alongside unequal crossing over in the rDNA, we believe there is little justification for assuming a random crossover point; thus, the first explanation merits further consideration.

Many advances have been made recently in understanding the biochemistry of recombination. In the best-studied system, the Chi system in prokaryotes (reviewed in Eggleston and West 1996), a double-strand break (DSB) is made, and the RecBCD complex unwinds and degrades the DNA until it encounters a Chi sequence from the 3′ side in the correct orientation. Further unwinding generates a single-strand tail with Chi at its 3′ end, which initiates pairing and strand exchange with the help of RecA. Thus, Chi has both orientation dependence and directionality. Holliday junctions are formed, and branch migration occurs via the action of RuvAB. Finally, resolution of the Holliday junction to generate recombinant molecules is catalyzed by RuvC, which preferentially nicks the DNA at a short target sequence. Thus, there are three potential sites that mediate recombination—initiation of DSB, initiation of strand exchange, and a signal to resolve the Holliday junction.

Biochemical understanding of eukaryote recombination lags behind that of prokaryotes, but many features seem to be conserved. If recombination in the Lp1 rDNA also required these three sites of recombination mediation, nonrandom crossover points could result. A potential model for bias of the crossover point to the left-hand side of the 111-/119-bp subrepeat array (in the orientation shown in Figure 6) is presented in Figure 7.

First, a DSB is made to the left of the 111-/119-bp subrepeat array. We have diagrammed this occurring in the 40-bp subrepeats, as Linares *et al.* (1994) in their examination of the unusual subrepeat organization in the *D. melanogaster* rDNA IGS implicated a smaller, more poorly maintained subrepeat array containing simple sequence motifs in the initiation of gene conversion that shapes the larger subrepeat array. An area of sequence simplicity has also been implicated in rDNA recombination in wheat (Barker *et al.* 1988). Therefore, subrepeat arrays with simple sequence motifs may act as recognition sites for the initiation of recombination, perhaps as the site of DSB.

The DSB may then be enlarged by exonuclease activity until an initiator of strand exchange with orientation dependence analogous to Chi is reached. It is interesting to note that the 111-bp subrepeat contains a sequence (AGTGGTGG; the reverse complement of the sequence shown in Figure 4) that is very similar to the Chi sequence (GCTGGTGG). This is in the orientation that would stimulate recombination if the DSB initiation point were to the left-hand side of the subrepeat array, and it contains the two paired guanosines that Hibner

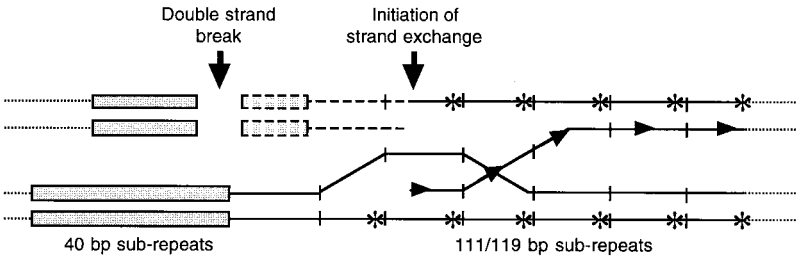

Figure 7.—Model detailing bias of the crossover point to the left-hand side of the 111-/119-bp subrepeat array. The diagram shows a hypothetical model for unequal crossing-over in the Lp1 IGS with the 111-/119-bp subrepeats misaligned. Initiation proceeds via a DSB (shown here in the 40-bp subrepeats) with exonuclease activity proceeding up to an initiator of strand exchange (shown as arrowheads in the 111-/119-bp subrepeats). The crossover point is then determined by the site of Holliday junction resolution. This preferentially occurs at a cleavage site (shown as asterisks) in the 111-/119-bp subrepeats once branch migration has proceeded to such a site. The 40-bp subrepeats are shown as shaded boxes. Individual 111-/119-bp subrepeats are delineated by short, vertical lines. Broken lines indicate exonuclease activity, and dotted lines indicate flanking sequence. The strand exchange initiator is only shown on the initiating strand, and the orientation dependence is indicated by the direction of the arrowheads. Cleavage of the Holliday junction on the outer strands as shown will result in reciprocal exchange products. The diagram is not drawn to scale.

et al. (1991) found to be present in many recombination-stimulating sequences.

Finally, branch migration would extend the resulting Holliday junction to a consensus site of a Holliday junction resolvase such as a topoisomerase I (Sekiguchi *et al.* 1996). If this site were in the 111-/119-bp subrepeats and occurred on the outer strands as diagrammed in Figure 7, then resolution would result in crossing over, with the crossover point biased to the left-hand side of the subrepeat array. The extent of branch migration would determine how far to the right the crossover point occurred. Bias of crossover to the left-hand side of the subrepeat array would conserve the subrepeat order at the right-hand end of the array while tending to rearrange the order at the left-hand end, unless misalignments occurred with the right-hand side of one of the subrepeat arrays. This is consistent with the observed results.

Lp1 is a hybrid organism, and the two progenitor isolates have been identified. The size of the "original" IGS in Lp1 is therefore known, as both progenitors have IGS lengths of ~4 kb. This leads to an interesting conclusion—the length heterogeneity in Lp1 is almost exclusively an increase in length, yet the reciprocal nature of unequal crossing over dictates that for every larger product formed, a smaller product must also be formed. The smallest IGS length is 3.5 kb, which would contain nine 111-/119-bp subrepeats—enough for more to be lost. This suggests a form of "selection" against short spacers. We propose some sort of homology interaction. It seems likely that the recombination equipment is limiting (Loidl and Nairz 1997), leading to competition between rDNA units for this equipment. Jinks-Robertson *et al.* (1993) and Yuan and Keil (1990) found that recombination in yeast between nontandem duplications requires at least 250 bp of homology, and the rate increases linearly up to 1 kb of homology, after which it plateaus off. Nine 111-/119-bp subrepeats represent ~1 kb; therefore efficient recombination might be achieved only by rDNA units with at least this number of subrepeats. It is interesting in this light that McKee *et al.* (1992) found that meiotic chromosome pairing in *D. melanogaster* required at least two rDNA IGS subrepeat elements. Spread of a repeat through an array by unequal crossing over requires that repeat to participate in unequal crossing over, so any repeats not participating in unequal crossing over, such as those with short 111-/119-bp subrepeat arrays, will be eliminated from the array as other repeats spread.

The IGS length heterogeneity arises through hybridization: The extraordinary length heterogeneity in the Lp1 IGS seems to be a consequence of the hybridization event, as neither of the Lp1 progenitors show such length heterogeneity. However, it does not seem to be an outcome of hybridization *per se*, as other hybrid endophytes from independent hybridization events do not display such IGS length heterogeneity (A. R. D. Ganley, unpublished results). Rather, control of length homogeneity seems to have been disrupted as a result of the hybridization. The nature of this disruption is not known, but could fall into three general categories: (1) loss of alignment control of the 111-/119-bp subrepeats, allowing misalignment that would generate the length heterogeneity and (2) loss of control of a maximum length for the 111-/119-bp subrepeat array. Stephan and Cho (1994) established a theoretical base for selection of a maximum size limit for a repeat array that is required for its maintenance, and Williams *et al.* (1987) invoked a similar form of selection on experimental grounds. Although no mechanism of action is known, it is possible that such an activity exists, or (3) alteration in the relative balance of gene conversion and unequal crossing over in favor of crossing over, destabilizing control of length. Such disruptions could be a consequence of chromosome expulsion during hybridization, gene doubling, or novel interactions between the two genomes analogous in a sense to the phenomenon of nucleolar dominance (Reeder and Roan 1984).

Concluding comments: The very nature of multigene families makes them recalcitrant to analysis of the mechanisms behind their evolution. We have demonstrated the occurrence of unequal crossing over in the rDNA

of Lp1, and this implies that unequal crossing over is a mechanism of homogenization in the concerted evolution of the rDNA. We have also presented circumstantial evidence for the specific initiation and for homology requirements of unequal crossing over, as well as circumstantial evidence for the involvement of another mechanism of homogenization, most probably gene conversion, in the concerted evolution of the rDNA in Lp1. As our understanding of the biochemistry and genetics of recombination increases, we will be able to transfer this knowledge to systems undergoing concerted evolution to test for similarities and differences. In the meantime, we must look to the systems that have been characterized, as well as to new systems, to find common threads that present testable ideas to tease out the details of the mechanisms responsible for homogenization.

We thank Carolyn Young, Kim Gan, Dianne Bird, and Kelly Marriott for expert technical assistance. We thank Mike Christensen for assistance with endophyte biology, especially the single-sporing work, and John Mackay for advice on PCR. We are also grateful to Max Scott, Brendon Monahan, and David Penny for helpful discussions. This work was supported by a Massey University Doctoral Scholarship to A.R.D.G.

LITERATURE CITED

Arnheim, N., and M. Kuehn, 1979 The genetic behavior of a cloned mouse ribosomal DNA segment mimics mouse ribosomal gene evolution. J. Mol. Biol. **134:** 743–765.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith *et al.*, 1987–1993 *Current Protocols in Molecular Biology*. John Wiley & Sons, New York.

Barker, R. F., N. P. Harberd, M. G. Jarvis and R. B. Flavell, 1988 Structure and evolution of the intergenic region in a ribosomal DNA repeat unit of wheat. J. Mol. Biol. **201:** 1–17.

Botchan, P., R. H. Reeder and I. B. Dawid, 1977 Restriction analysis of the nontranscribed spacers of *Xenopus laevis* ribosomal DNA. Cell **11:** 599–607.

Brown, D. D., P. C. Wensink and E. Jordan, 1972 A comparison of the ribosomal DNA's of *Xenopus laevis* and *Xenopus mulleri*: the evolution of tandem genes. J. Mol. Biol. **63:** 57–73.

Brownlee, A. G., 1988 A rapid DNA isolation procedure applicable to many refractory filamentous fungi. Fungal Genet. Newsl. **35:** 8–9.

Bullock, W. O., J. M. Fernandez and J. M. Short, 1987 XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. Biotechniques **5:** 376–378.

Christensen, M. J., A. Leuchtmann, D. D. Rowan and B. A. Tapper, 1993 Taxonomy of Acremonium endophytes of tall fescue (*Festuca arundinacea*), meadow fescue (*F. pratensis*) and perennial ryegrass (*Lolium perenne*). Mycol. Res. **97:** 1083–1092.

Coen, E. S., J. M. Thoday and G. Dover, 1982 Rate of turnover of structural variants in the rDNA gene family of *Drosophila melanogaster*. Nature **295:** 564–568.

Collett, M. A., R. E. Bradshaw and D. B. Scott, 1995 A mutualistic fungal symbiont of perennial ryegrass contains two different *pyr4* genes, both expressing orotidine-5′-monophosphate decarboxylase. Gene **158:** 31–39.

Copenhaver, G. P., and C. S. Pikaard, 1996 Two-dimensional RFLP analyses reveal megabase-sized clusters of rRNA gene variants in *Arabidopsis thaliana*, suggesting local spreading of variants as the mode for gene homogenization during concerted evolution. Plant J. **9:** 273–282.

Crease, T. J., 1995 Ribosomal DNA evolution at the population level: nucleotide variation in intergenic spacer arrays of *Daphnia pulex*. Genetics **141:** 1327–1337.

Dobson, J. M., 1997 The molecular characterization of the HMG CoA reductase gene from *Neotyphodium lolii*. M.Sc. Thesis, Massey University, Palmerston North, New Zealand.

Dover, G. A., 1982 Molecular drive: a cohesive mode of species evolution. Nature **299:** 111–117.

Dover, G. A., A. R. Linares, T. Bowen and J. M. Hancock, 1993 Detection and quantification of concerted evolution and molecular drive. Methods Enzymol. **224:** 525–541.

Dvorák, J., and R. Appels, 1986 Investigation of homologous crossing over and sister chromatid exchange in the wheat *Nor-B2* locus coding for rDNA and *Gli-B2* locus coding for gliadins. Genetics **113:** 1037–1056.

Dvorák, J., D. Jue and M. Lassner, 1987 Homogenization of tandemly repeated nucleotide sequences by distant-dependent nucleotide sequence conversions. Genetics **116:** 487–498.

Edelman, G. M., and J. A. Gally, 1970 Arrangement and evolution of eukaryotic genes, pp. 962–972 in *The Neurosciences: Second Study Program*, edited by F. O. Schmitt. Rockefeller University Press, New York.

Eggleston, A. K., and S. C. West, 1996 Exchanging partners: recombination in *E. coli*. Trends Genet. **12:** 20–26.

Elder, J. F., Jr., and B. J. Turner, 1995 Concerted evolution of repetitive DNA sequences in eukaryotes. Q. Rev. Biol. **70:** 297–320.

Feinberg, A. P., and B. Vogelstein, 1983 A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:** 6–13.

Frischauf, A.-M., H. Lehrach, A. Poustka and N. Murray, 1983 Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170:** 827–842.

Hibner, B. L., W. D. Burke and T. H. Eickbush, 1991 Sequence identity in an early chorion multigene family is the result of localized gene conversion. Genetics **128:** 595–606.

Hillis, D. M., C. Moritz, C. A. Porter and R. J. Baker, 1991 Evidence for biased gene conversion in concerted evolution of ribosomal DNA. Science **252:** 308–310.

Holliday, R., 1964 A mechanism for gene conversion in fungi. Genet. Res. **5:** 282–304.

Israelewski, N., and E. R. Schmidt, 1982 Spacer size heterogeneity in ribosomal DNA of *Chironomus thummi* is due to a 120 bp repeat homologous to a predominantly centromeric repeated sequence. Nucleic Acids Res. **10:** 7689–7700.

Jackson, J. A., and G. R. Fink, 1981 Gene conversion between duplicated genetic elements in yeast. Nature **292:** 306–311.

Jeffreys, A. J., A. MacLeod, K. Tamaki, D. L. Neil and D. G. Monckton, 1991 Minisatellite repeat coding as a digital approach to DNA typing. Nature **354:** 204–209.

Jinks-Robertson, S., and T. D. Petes, 1993 Experimental determination of rates of concerted evolution. Methods Enzymol. **224:** 631–646.

Jinks-Robertson, S., M. Michelitch and S. Ramcharan, 1993 Substrate length requirements for efficient mitotic recombination in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **13:** 3937–3950.

Kellogg, E. A., and R. Appels, 1995 Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. Genetics **140:** 325–343.

Kelly, R. J., R. D. Johnson and A. Siegel, 1990 Heterogeneity and organization of the ribosomal RNA genes of *Cucurbita maxima*. Plant Mol. Biol. **14:** 927–933.

Krystal, M., and N. Arnheim, 1978 Length heterogeneity in a region of the human ribosomal gene spacer is not accompanied by extensive population polymorphism. J. Mol. Biol. **126:** 91–104.

Li, W.-H., 1997 *Molecular Evolution*. Sinauer Associates, Sunderland, MA.

Linares, A. R., T. Bowen and G. A. Dover, 1994 Aspects of nonrandom turnover involved in the concerted evolution of intergenic spacers within the ribosomal DNA of *Drosophila melanogaster*. J. Mol. Evol. **39:** 151–159.

Loidl, J., and K. Nairz, 1997 Karyotype variability in yeast caused by nonallelic recombination in haploid meiosis. Genetics **146:** 79–88.

Long, E. O., and I. B. Dawid, 1980 Repeated genes in eukaryotes. Annu. Rev. Biochem. **49:** 727–764.

Martin, F. N., 1990 Variation in the ribosomal DNA repeat unit within single-oospore isolates of the genus Pythium. Genome **33:** 585–591.

McKee, B. D., L. Habera and J. A. Vrana, 1992 Evidence that intergenic spacer repeats of *Drosophila melanogaster* rRNA genes function as X-Y pairing sites in male meiosis, and a general model for achiasmatic pairing. Genetics **132:** 529–544.

Meselson, M. S., and C. M. Radding, 1975 A general model for genetic recombination. Proc. Natl. Acad. Sci. USA **72:** 358–361.

Nagylaki, T., and T. D. Petes, 1982 Intrachromosomal gene conversion and the maintenance of sequence homogeneity among repeated genes. Genetics **100:** 315–337.

Ohta, T., 1976 Simple model for treating evolution of multigene families. Nature **263:** 74–76.

Orr-Weaver, T. L., and J. W. Szostak, 1985 Fungal recombination. Microbiol. Rev. **49:** 33–58.

Petes, T. D., 1980 Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes. Cell **19:** 765–774.

Petes, T. D., and D. Botstein, 1977 Simple Mendelian inheritance of the reiterated ribosomal DNA of yeast. Proc. Natl. Acad. Sci. USA **74:** 5091–5095.

Reeder, R. H., and J. G. Roan, 1984 The mechanism of nucleolar dominance in Xenopus hybrids. Cell **38:** 39–44.

Reeder, R. H., D. D. Brown, P. K. Wellauer and I. B. Dawid, 1976 Patterns of ribosomal DNA spacer lengths are inherited. J. Mol. Biol. **105:** 507–516.

Rogers, S. O., and A. J. Bendich, 1987 Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. Plant Mol. Biol. **9:** 509–520.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Samson, M.-L., and M. Wegnez, 1988 Bipartite structure of the 5S ribosomal gene family in a *Drosophila melanogaster* strain, and its evolutionary implications. Genetics **118:** 685–691.

Schafer, M., A. R. Wyman and R. White, 1981 Length variation in the non-transcribed spacer of *Calliphora erythrocephala* ribosomal DNA is due to a 350 base-pair repeat. J. Mol. Biol. **146:** 179–199.

Schardl, C. L., J.-S. Liu, J. F. White, Jr., R. A. Finkel, Z. An *et al.*, 1991 Molecular phylogenetic relationships of nonpathogenic grass mycosymbionts and clavicipitaceous plant pathogens. Plant. Syst. Evol. **178:** 27–41.

Schardl, C. L., A. Leuchtmann, H.-F. Tsai, M. A. Collett, D. M. Watt *et al.*, 1994 Origin of a fungal symbiont of perennial ryegrass by interspecific hybridization of a mutualist with the ryegrass choke pathogen, *Epichloë typhina*. Genetics **136:** 1307–1317.

Schlötterer, C., and D. Tautz, 1994 Chromosomal homogeneity of Drosophila ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. Curr. Biol. **4:** 777–783.

Sekiguchi, J., N. C. Seeman and S. Shuman, 1996 Resolution of Holliday junctions by eukaryotic DNA topoisomerase I. Proc. Natl. Acad. Sci. USA **93:** 785–789.

Smith, G. P., 1973 Unequal crossing over and the evolution of multigene families. Cold Spring Harbor Symp. Quant. Biol. **38:** 507–513.

Stephan, W., and S. Cho, 1994 Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. Genetics **136:** 333–341.

Szostak, J. W., and R. Wu, 1980 Unequal crossing over in the ribosomal DNA of *Saccharomyces cerevisiae*. Nature **284:** 426–430.

Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein and F. W. Stahl, 1983 The double-strand-break repair model for recombination. Cell **33:** 25–35.

Tautz, D., C. Tautz, D. Webb and G. A. Dover, 1987 Evolutionary divergence of promoters and spacers in the rDNA family of four Drosophila species: implications for molecular coevolution in multigene families. J. Mol. Biol. **195:** 525–542.

Toda, T., Y. Nakaseko, O. Niwa and M. Yanagida, 1984 Mapping of rRNA genes by integration of hybrid plasmids in *Schizosaccharomyces pombe*. Curr. Genet. **8:** 93–97.

Tsai, H.-F., J.-S. Liu, C. Staben, M. J. Christensen, G. C. M. Latch *et al.*, 1994 Evolutionary diversification of fungal endophytes of tall fescue grass by hybridization with Epichloë species. Proc. Natl. Acad. Sci. USA **91:** 2542–2546.

Vieira, J., and J. Messing, 1987 Production of single-stranded plasmid DNA. Methods Enzymol. **153:** 3–11.

Wellauer, P. K., and I. B. Dawid, 1978 Ribosomal DNA in *Drosophila melanogaster*: heteroduplex mapping of cloned and uncloned rDNA. J. Mol. Biol. **126:** 769–782.

Wellauer, P. K., I. B. Dawid, D. D. Brown and R. H. Reeder, 1976 The molecular basis for length heterogeneity in ribosomal DNA from *Xenopus laevis*. J. Mol. Biol. **105:** 461–486.

Wendel, J. F., A. Schnabel and T. Seelanan, 1995 Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (Gossypium). Proc. Natl. Acad. Sci. USA **92:** 280–284.

White, T. J., T. Bruns, S. Lee and J. Taylor, 1990 Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, pp. 315–322 in *PCR Protocols: A Guide to Methods and Applications*, edited by M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White. Academic Press, San Diego.

Whitehouse, H. L. K., 1982 *Genetic Recombination, Understanding the Mechanisms*. John Wiley & Sons, New York.

Williams, S. M., G. R. Furnier, E. Fuog and C. Strobeck, 1987 Evolution of the ribosomal DNA spacers of *Drosophila melanogaster*: different patterns of variation on *X* and *Y* chromosomes. Genetics **116:** 225–232.

Williams, S. M., J. A. Kennison, L. G. Robbins and C. Strobeck, 1989 Reciprocal recombination and the evolution of the ribosomal gene family of *Drosophila melanogaster*. Genetics **122:** 617–624.

Yuan, L.-W., and R. L. Keil, 1990 Distance independence of mitotic intrachromosomal recombination in *Saccharomyces cerevisiae*. Genetics **124:** 263–273.

Communicating editor: G. B. Golding