Multiple Recombination Events Maintain Sequence Identity Among Members of the Nitrogenase Multigene Family in *Rhizobium etli*

César Rodríguez and David Romero

Departamento de Genética Molecular, Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

Manuscript received December 3, 1997
Accepted for publication March 2, 1998

ABSTRACT

A distinctive characteristic of the Rhizobium genome is the frequent finding of reiterated sequences, which often constitute multigene families. Interestingly, these families usually maintain a high degree of nucleotide sequence identity. It is commonly assumed that apparent gene conversion between reiterated elements might lead to concerted variation among members of a multigene family. However, the operation of this mechanism has not yet been demonstrated in the Rhizobiaceae. In this work, we employed different genetic constructions to address the role of apparent gene conversion as a homogenizing mechanism between members of the plasmid-located nitrogenase multigene family in *Rhizobium etli*. Our results show that a 28-bp insertion into one of the nitrogenase reiterations can be corrected by multiple recombination events, including apparent gene conversion. The correction process was dependent on the presence of both a wild-type *recA* gene and wild-type copies of the nitrogenase reiterations. Frequencies of apparent gene conversion to the wild-type nitrogenase reiterations were the same when the insertion to be corrected was located either in *cis* or in *trans*, indicating that this event frequently occurs through intermolecular interactions. Interestingly, a high frequency of multiple crossovers was observed, suggesting that these large plasmid molecules are engaging repeatedly in recombination events, in a situation akin to phage recombination or recombination among small, high-copy number plasmids.

GENE duplication, leading to the formation of multigene families in eukaryotes, has been widely recognized as a mechanism for the generation of new functions (Ohno 1970; John and Miklos 1988). Interestingly, members of multigene families tend to vary in a concerted way, keeping a high similarity between their members (John and Miklos 1988; Dover 1993). Concerted evolution between members of tandemly-arranged multigene families in eukaryotes has been amply documented. Conservation in a nucleotide sequence is thought to occur predominantly through frequent unequal exchanges between its members (Petes 1980; Szostak and Wu 1980; Williams and Strobeck 1985). A consequence of this mechanism is the frequent expansion and contraction of the tandem array. However, recent determinations show that an alternate mechanism, gene conversion, plays a major role in achieving homogenization in tandem multigene families (Gangloff *et al.* 1996). Gene conversion has been defined as the non-reciprocal transfer of sequence information between homologous or homeologous DNA sequences. Frequent events of gene conversion are also responsible for concerted evolution between members of dispersed gene families (Jackson and Fink 1981; Klein and Petes 1981). Unequal exchanges are less adequate to achieve homogenization in these cases, due to the high likelihood of rearrangement of single-copy DNA flanked by dispersed members.

Studies about the occurrence of similar processes in prokaryotes have been hindered by the paucity of reiterated elements in enterobacterial genomes. A recent determination, based on the sequence of the whole *Escherichia coli* genome (Blattner *et al.* 1997), shows that reiterated elements constitute about 2.5% of the genome. The most conspicuous families of long reiterated elements in *E. coli* and *Salmonella typhimurium* are the *rrn* operons, *tuf* genes, and different types of insertion sequences (Bachellier *et al.* 1996; Deonier 1996; Blattner *et al.* 1997), which are commonly arranged as dispersed reiterations. Typically, a high level of nucleotide sequence identity is observed among members of each family. Several reports indicate that homogenization between the reiterated *rrn* operons in *E. coli* (Harvey and Hill 1990), duplicated flagellin genes (Okazaki *et al.* 1993) or the *tuf* reiterations in *S. typhimurium* (Abdulkarim and Hughes 1996) may be achieved through apparent gene conversion.

Extensive DNA reiteration is found in the genomes of bacteria belonging to the symbiotic nitrogen-fixing genus Rhizobium (reviewed by Romero *et al.* 1997). These genomes may carry as much as 700 representatives of long, reiterated elements, belonging to 200 different families (Flores *et al.* 1987). This high level of

Corresponding author: David Romero, Depto. de Genética Molecular, Centro de Investigación sobre Fijación de Nitrógeno-UNAM, Apartado Postal 565-A, 62100 Cuernavaca, Morelos, México.
E-mail: dromero@cifn.unam.mx

reiteration has been observed both in the chromosome and in the large plasmids that are typical of this genus (Flores *et al.* 1987; Girard *et al.* 1991). In fact, an analysis of the sequence of the whole symbiotic plasmid (pSym, 536.1 kb) of *Rhizobium* sp. NGR234 reveals that reiterated elements constitute about 18% of the plasmid genome (Freiberg *et al.* 1997).

Besides transposable elements, multigenic families have been observed for several housekeeping genes, such as the *fla* (Bergman *et al.* 1991), *ftsZ* (Margolin and Long 1994), *groEL* (Fischer *et al.* 1993), *rpoN* (Kündig *et al.* 1993) and citrate synthase genes (Pardo *et al.* 1994; Hernández-Lucas *et al.* 1995). Several genes involved solely in the symbiotic process are also reiterated, such as *nifHDK* (Badenoch-Jones *et al.* 1989; Norel and Elmerich 1987; Quinto *et al.* 1985), *fixN* (David *et al.* 1987; Schlüter *et al.* 1997) and the *nodD* genes (Schultze *et al.* 1994). The reiteration mode is usually dispersed; only the *fla* gene reiterations are arranged in tandem (Bergman *et al.* 1991).

High levels of nucleotide identity are common between these reiterations. A considerable fraction of the reiterated class (about 70%) in the pSym of *Rhizobium* sp. NGR234 is comprised of identical reiterations (Freiberg *et al.* 1997). Other multigenic families, such as *fla* (Bergman *et al.* 1991), *nifHDK* (Badenoch-Jones *et al.* 1989; Norel and Elmerich 1987; Quinto *et al.* 1985) and the citrate synthase genes (Pardo *et al.* 1994; Hernández-Lucas *et al.* 1995) also exhibit high identities (over 95%) in nucleotide sequence. A trivial explanation for the occurrence of such high levels of identity would be a relatively recent evolutionary origin. This explanation seems unlikely, because (i) in some cases, such as the *nifHDK* family of *Rhizobium etli*, the sequences are ancient enough to be present in every member of the species (Segovia *et al.* 1993), and (ii) although high levels of nucleotide sequence identity are seen between reiterations within a species, a lower level is seen between species themselves. These data have led to the suggestion that some homogenizing mechanism, conceivably gene conversion, is operating between reiterated sequences in Rhizobium.

So far, no studies have been published about the occurrence of gene conversion-like events between reiterations in Rhizobium. The only data concerning this phenomenon were obtained during a study of phage crosses in *Rhizobium meliloti*, where apparent gene conversion is claimed to occur at a low frequency (Orosz *et al.* 1980), close to the one for spontaneous mutation. If the low frequency observed also applies to recombinational interactions between reiterations, apparent gene conversion would be an inefficient homogenizing mechanism in Rhizobium.

In this work, we address the role of apparent gene conversion as an homogenizing mechanism between members of the nitrogenase multigene family in *R. etli*. Our results indicate that multiple recombination events, including apparent gene conversion, play an important role in maintaining sequence identity among members of this family.

MATERIALS AND METHODS

Bacterial strains, plasmids and media: The bacterial strains and plasmids employed are listed in Table 1. *E. coli* strains were grown at 37° in luria broth medium (Miller 1972), and *R. etli* strains were grown at 30° in PY medium (Noel *et al.* 1984). Antibiotics were added at the following concentrations: kanamycin (Km), 30 μg ml^{-1} (*E. coli*) or at variable concentrations depending on the purpose (*R. etli*, see below); nalidixic acid (Nal), 20 μg ml^{-1} (*R. etli*); spectinomycin (Sp), 100 μg ml^{-1} (*E. coli* or *R. etli*); and tetracycline (Tc), 5 to 10 μg ml^{-1} (*R. etli*) or 10 μg ml^{-1} (*E. coli*).

Plasmid construction: All DNA manipulations were carried out under standard protocols (Sambrook *et al.* 1989) using restriction enzymes, bacterial alkaline phosphatase and T4 polynucleotide ligase (Amersham Ltd., England). To construct a hybrid Km^r/Sp^r cassette, a 1074-bp fragment, containing the Sp^r gene from pHP45-ΩSp (Table 1) was obtained through PCR amplification employing oligonucleotides o4 and o5 (Table 2); these oligonucleotides contain *Bgl*II sites. This PCR product was digested with *Bgl*II and ligated into *Bam*HI-restricted pSUP5011 (a plasmid containing Tn*5-mob*, Simon 1984), replacing the 1.6-kb *mob* fragment. The resulting plasmid, which carries a Tn*5* Km/Sp, was called pCRS1. From pCRS1, digestion of the *Bgl*II sites already existing in the IS*50*s led to the release of a 3814-bp *Bgl*II fragment, containing a promoterless Km^r gene and the Sp^r gene with its own promoter.

To facilitate further subcloning steps of this fragment, the polylinker-containing vector pIC20H (Marsh *et al.* 1984) was digested with *Sac*I and religated, removing a segment carrying *Xho*I, *Bgl*II, *Xba*I, *Eco*RV, *Cla*I and *Eco*RI sites from the polylinker; this plasmid was called pCRS2. The 3814-bp *Bgl*II fragment from pCRS1 was ligated into the unique *Bam*HI site of pCRS2, generating pCRS3. From pCRS3, the promoterless Km^r gene and the Sp^r gene with its own promoter were excised as a 3864-bp *Hin*dIII fragment, through digestion of two *Hin*dIII sites on the polylinker.

To generate a transcriptional fusion between this fragment and the *nifD* gene from *R. etli*, plasmid pEM15 (Morett *et al.* 1988) was used. This plasmid is a derivative of pSUP205 (Simon *et al.* 1983), containing one of the *nifHDK* operons of *R. etli* as an *Eco*RI fragment (*nifHDK* region a). Ligation of the 3864 bp *Hin*dIII fragment from pCRS3 into one of the *Hin*dIII sites present in pEM15 provokes an interruption of the *nifD* coding sequence (codon 139) creating, in the proper orientation, a transcriptional fusion between *nifD* and the promoterless Km^r gene. The resulting plasmid was called pCRS4.

To generate a polar insertion in *nifH*, a 28-bp double-stranded oligonucleotide (I-SceI, Table 2) with overhanging, compatible *Bgl*II ends, was ligated into the unique *Bgl*II site of pCRS4, thus interrupting the *nifH* coding sequence (codon 147); this plasmid was named as pCRS5. Finally, to generate plasmid derivatives able to replicate in *R. etli*, plasmids pCRS4 and pCRS5 were digested with *Eco*RI and the resulting fragments were ligated separately onto the unique *Eco*RI site of the broad-host range plasmid pRK7813 (Jones and Gutterson 1987); the resulting plasmids were called pCRS6 and pCRS7, respectively.

Construction of *Rhizobium etli* strains carrying the *nifH::I-SceI* and *nifD::*Km/Sp alleles: Introduction of the *nifH::I-SceI* and *nifD::*Km/Sp alleles into *R. etli* was carried out by an *in*

TABLE 1

Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant properties	Source or reference
Rhizobium etli		
CFN42	Wild-type strain	Quinto *et al.* (1985)
CE3	Sm^r derivative from CFN42	Noel *et al.* (1984)
CFNX89	CE3 derivative lacking the pSym	Brom *et al.* (1992)
CFNX236	CE3 derivative carrying a *nifD*::Km/Sp allele on the pSym	This study
CFNX237	CE3 derivative carrying the *nifH::I-SceI* and *nifD*::Km/Sp allelic combination on the pSym	This study
CFNX238	CE3 derivative carrying plasmid pCRS6	This study
CFNX239	CE3 derivative carrying plasmid pCRS7	This study
CFNX240	CFNX89 derivative carrying plasmid pCRS6	This study
CFNX241	CFNX89 derivative carrying plasmid pCRS7	This study
CFNX242	A *recA*::ΩCm derivative from strain CFNX237	This study
Escherichia coli		
HB101	Host for recombinant plasmids	Boyer and Roulland-Dussoix (1969)
S17-1	C600::RP-4-2 (Tc::Mu) (Km::Tn7)	Simon (1984)
Plasmids		
pEM15	A pSUP205 derivative carrying *nifHDK* region a from *R. etli* CFN42	Morett *et al.* (1988)
pHP45ΩSp	Source of the ΩSp cassette	Prentki and Krisch (1984)
pIC20H	Polylinker-containing vector	Marsh *et al.* (1984)
pMC71A	Plasmid carrying a constitutive *nifA* gene from *Klebsiella pneumoniae*, Cm^r	Buchanan-Wollaston *et al.* (1981)
pMS22	Suicide vector carrying the *recA*::ΩCm allele from *R. etli*	J. Martínez-Salazar, unpublished results
pRK2013	Conjugation helper	Figurski and Helinski (1979)
pRK7813	Broad-host range plasmid, Tc^r	Jones and Gutterson (1987)
pSUP205	Suicide vector, Cm^r, Tc^r	Simon *et al.* (1983)
pSUP5011	A pSUP205 derivative carrying Tn*5-mob*	Simon (1984)
pCRS1	Intermediate plasmid carrying a hybrid Tn*5* Km/Sp	This study
pCRS2	Derivative from pIC20H with a deletion in the polylinker	This study
pCRS3	Derivative from pCRS2 containing a promoterless Km^r gene and the Sp^r gene with its own promoter	This study
pCRS4	Derivative from pEM15 carrying the *nifD*::Km/Sp allele	This study
pCRS5	Derivative from pCRS4 carrying the *nifH::I-SceI* and *nifD*::Km/Sp allelic combination	This study
pCRS6	Derivative from pRK7813 carrying the *nifD*::Km/Sp allele	This study
pCRS7	Derivative from pRK7813 carrying the *nifH::I-SceI* and *nifD*::Km/Sp allelic combination	This study

vivo gene replacement procedure (Simon *et al.* 1983). To that end, plasmid pCRS4 was introduced by transformation into *E. coli* S17-1 and the transformants were mated with *R. etli* CE3 as a recipient. Double recombinants were selected as Nal^r Sp^r Tc^s transconjugants. To verify that the desired gene replacement has occurred, double recombinants were analyzed by Southern blot hybridization against the appropriate *nif* and ΩSp probes. This procedure yielded strain CFNX236, which carries the *nifD*::Km/Sp allele in *nif* region a. Introduction of the *nifH::I-SceI* and *nifD*::Km/Sp allelic combination was carried out in the same way, but employing pCRS5 as the donor. To ensure coinheritance of both markers, double recombinants were analyzed by PCR amplification using oligonucleotides o2 and o3 (see Table 2 and Figure 1); this primer pair only yielded a PCR product upon integration of the *nifH::I-SceI* allele. A strain containing both the *nifH::I-SceI* and *nifD*::Km/Sp allelic combination in *nif* region a was called CFNX237. A *recA*::ΩCm derivative from strain CFNX237 was generated by a homogenotization procedure developed in our laboratory (Martínez-Salazar *et al.* 1991; Romero *et al.* 1995), using pMS22 as the source of the *recA* allele. This procedure gave strain CFNX242.

To generate *R. etli* derivatives carrying these allelic combinations onto small, self-replicating plasmids, *E. coli* S17-1 derivatives carrying pCRS6 or pCRS7 were mated with either *R. etli* CE3 (pSym$^+$) or *R. etli* CFNX89 (pSym$^-$) as recipients. Transconjugants were selected by their resistance to both nalidixic acid and spectinomycin. These crosses produced strains CFNX238 to CFNX241 (see Table 1).

PCR amplification and nucleotide sequencing: PCR amplifications were carried out using AmpliTaq DNA polymerase in a DNA Thermal Cycler 480 (Perkin-Elmer, Inc., Norwalk, CT). PCR conditions consisted of 30 cycles of 92° for 1 min, 56° for 1 min, and 72° for 1 min, except for amplifications

TABLE 2

Oligonucleotides used in this study

Name	Sequence	Source: complementary nucleotides	Accession no.
o1	5′-GAGGACCTGCTCAAGGCCGGCTAC-3′	*nifH*: 423-446	M10587
o2	5′-CGCTAGGGATAACAGGGTAATATA-3′	I-SceI: 1-24	See below
o3	5′-CATCTTCCTGAGCTCGGCGTGCTG-3′	*nifH*: 878-855	M10587
o4	5′-GAAGATCTCCTGATAGTTTGGCTGTGAG-3′	ΩSp: 1784-1756	M60473
o5	5′-TAAGATCTCAGTGGCGGTTTTCATGGCT- 3′	ΩSp: 712-740	M60473
I-SceI	5′-GATCCGCTAGGGATAACAGGGTAATATA-3′ 3′-GCGATCCCTATTGTCCCATTATATCTAG-5′	Not applicable[a]	Not applicable[a]

Nucleotides corresponding to the *Bgl* II site in o4 and o5 are underlined.
[a] Commercial oligonucleotide (Boehringer-Mannheim, Germany).

employing primer pairs o1–o4 and o2–o4, that were done by 30 cycles of 95° for 1 min, 55° for 1 min, and 72° for 2 min. PCR products for nucleotide sequencing were purified using Centri-Sep spin columns (Princeton Separations Inc., Adelphia, NJ). Nucleotide sequencing was performed with an Applied Biosystems Inc. model 373A automated DNA sequencer and a Taq DyeDeoxy Terminator cycle sequencing kit as specified by the manufacturer (Applied Biosystems Inc., Foster City, CA).

Filter blot hybridization and determination of plasmid profiles: Genomic DNA was digested with *Bam*HI, electrophoresed in 1% agarose gels, blotted onto nitrocellulose (Hybond N+), and hybridized under stringent conditions using Amersham's Rapid-hyb buffer as specified by the manufacturer (Amersham Corp.). Plasmid profiles were obtained by an in-gel lysis method (Eckhardt 1978), blotted onto nitrocellulose and hybridized similarly. Hybridizations with oligonucleotide probes were done in a sodium chloride-sodium citrate solution using standard procedures (Ausubel *et al.* 1987). Most probes were linearized and labelled with ^{32}P-α-CTP by a random priming procedure (Feinberg and Vogelstein 1983) using a Rediprime DNA labelling system (Amersham Corp.). Oligonucleotide probes were labelled with ^{32}P-γ-ATP using T4 polynucleotide kinase.

Molecular characterization of Km^r derivatives: To ascertain the molecular events leading to the formation of each Km^r derivative, single-colony isolates were initially screened for the presence and location of the *nifH::I-SceI* allele. To that end, genomic DNA of each isolate was subjected to PCR employing either the o2–o3 primer pair or the o1–o3 pair (Figure 1). Primer o2 has a sequence that matches the *nifH::I-SceI* allele, while primers o1 and o3 bind to specific points in the *nifH* sequence (Table 2). Thus, reactions with the o2–o3 pair gave an amplified product only if the *nifH::I-SceI* allele was still present, while those with pair o1–o3 served as a positive control for the PCR. To determine the location of the *nifH::I-SceI* allele in the Km^r derivatives that still carried this allele, further PCR amplifications were made with primer pairs o2–o4 and o1–o4 (Figure 1 and Table 2). The first pair of primers gave a 3665-bp PCR product only if the *nifH::I-SceI* allele was still coupled to the *nifD*::Km/Sp allele, while the second pair provided a positive control for these reactions.

Km^r derivatives were also characterized by determining pSym size in Eckhardt-type gels. This analysis allows us to distinguish wild-type plasmids (390 kb) from amplified (510 kb) or deleted (270 kb) derivatives, as described previously (Romero *et al.* 1991; Romero *et al.* 1995). Additionally, genomic DNA of each derivative was digested with *Bam*HI, blotted, and probed against a *nifH*-specific probe. Under these conditions, strain CFNX237 shows three hybridizing bands of 13 kb (*nif* region a harboring the *nifH::I-SceI* and *nifD::*Km/Sp allelic combination), 5.6 kb (*nif* region b) and 4.5 kb (*nif* region c). Derivatives harboring a tandem amplification preserve the same three bands, but show an additional 8.8-kb band, representing the join point; stoichiometry of these bands is also typical, where the band corresponding to *nif* region c is more abundant than the rest. Band pattern is also altered in the derivative carrying an inversion, showing three bands of 10.4, 6.5 and 5.6 kb; the first two bands are join points for this rearrangement, while the last corresponds to *nif* region b. The derivative carrying a deletion show a single, *nifH*-positive band of 13.6 kb; this band is the join point. Location of the *nifH::I-SceI* and *nifD::*Km/Sp alleles was verified through hybridization with allele-specific probes.

RESULTS

Experimental design: The main objective of this work was to study the relative role of recombination *vs.* apparent gene conversion in the maintenance of sequence identity in a reiterated multigene family. To that end, we chose the nitrogenase multigene family of *R. etli* as a model. All members of this family are located in a single 390-kb plasmid, the symbiotic plasmid (pSym). This family is composed of two identical direct reiterations of about 5 kb (*nif* regions a and b), which are *nifHDK* operons; these operons are 120 kb apart on the pSym. The third element of this family (*nif* region c) is located in the middle of this zone, and consists of an identical reiteration, 1.5 kb long, harboring a complete *nifH* gene and a truncated *nifD* gene in an inverted orientation *vis a vis nif* regions a and b. Homologous recombination between *nif* regions a and b leads to frequent genomic rearrangements, such as deletions and amplifications (Romero *et al.* 1991; Romero *et al.* 1995). This particular arrangement allows us to study the recombinational dynamics of a plasmidic multigene family.

Since expression of these *nif* regions under *ex planta* conditions does not confer any scorable phenotype, we modified *nif* region a by inserting a promoterless Km^r cassette into the *nifD* gene, as described in materials

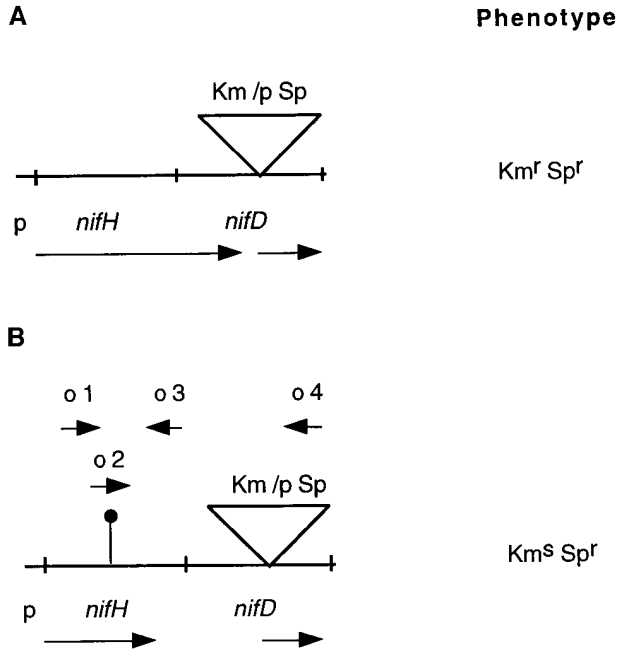

Figure 1.—Experimental design. Both parts of the figure represent *nif* region a. Symbols are as follows: the *nifH* promoter (p); the *nifD*::Km/Sp allele (triangle symbol); and the *nifH*::*I-SceI* allele (stick and ball symbol). Arrows beneath the figures indicate the expected transcripts in each case. (a) In this case, cells become Km^r due to transcription of the Km^r gene from the *nifH* promoter. (b) Introduction of the *nifH*::*I-SceI* allele leads to a Km^s phenotype, due to the introduction of additional termination codons. Small arrows in this figure indicate the location of specific oligonucleotides (o1–o4) used for characterization.

and methods. This cassette also carries a Sp^r gene with its own promoter. As shown in Figure 1a, expression of the Km^r gene in this *nifD*::Km/Sp allele should be under the control of the *nifH* promoter. This construct was then modified by the insertion of a 28-bp oligonucleotide into the *nifH* gene (the *nifH*::*I-SceI* allele, see materials and methods). This *nifH*::*I-SceI* allele leads to alterations in the translational reading of the *nifH* gene, because in-frame reading of this insertion causes misreading of two termination codons (UAG and UAA) present in the oligonucleotide. Additionally, since this insertion provokes a +1 frameshift, two additional stop codons are uncovered (UAA and UGA) at positions matching codons 185–186 and 214–215 of the wild-type *nifH* sequence, respectively. As shown in Figure 1b, the *nifH*::*I-SceI* allele should block, by polarity, the expression of the *nifD*::Km/Sp allele, thus leading to a Km^s phenotype. Selection for Km^r derivatives give us a positive system to identify events that lead to the loss or relocation of the 28-bp insertion, conceivably *via* recombination with the other members of this multigene family.

Initial tests of the functionality of this system were done in *E. coli*. To that end, we introduced plasmids pCRS4 (carrying the *nifD*::Km/Sp allele; Table 1) and pCRS5 (carrying the *nifH*::*I-SceI* and *nifD*::Km/Sp allelic combination; Table 1) into *E. coli* HB101. In this system, expression directed by the *nifH* promoter depends on σ^{54} and the NifA activator protein (Valderrama *et al.* 1996). As expected, *E. coli* strains harboring either pCRS4 or pCRS5 were sensitive to kanamycin (60 μg ml^{-1}). Upon introduction of a second, compatible plasmid carrying the constitutively-activated *Klebsiella pneumoniae nifA* gene (pMC71A) into the strain harboring pCRS4, cells became resistant to kanamycin. In contrast, cells carrying both pCRS5 and pMC71A remained sensitive to kanamycin. Thus, these results indicate that expression of the Km^r gene in the *nifD*::Km/Sp allele is dependent on the *nifH* promoter and that the introduction of the *nifH*::*I-SceI* allele blocks that expression.

Construction of *R. etli* strains containing on the pSym either the *nifD*::Km/Sp allele (strain CFNX236, Table 1) or the *nifH*::*I-SceI* and *nifD*::Km/Sp allelic combination (strain CFNX237, Table 1) was done by allelic replacement (see materials and methods). In this host, maximal expression from the *nifH* promoter is achieved under microaerobic conditions (Valderrama *et al.* 1996). However, basal transcription from this promoter under aerobic conditions was enough to confer to strain CFNX236 a low-level resistance to kanamycin (3 μg ml^{-1}). As expected, strain CFNX237, carrying the *nifH*::*I-SceI* and *nifD*::Km/Sp allelic combination was sensitive to kanamycin. Thus, loss or relocation of the *nifH*::*I-SceI* allele can be detected in strain CFNX237 by scoring the frequency of Km^r derivatives.

Theoretically, recombinational repair of the *nifH*::*I-SceI* allele to yield a Km^r derivative can arise either by sister-strand exchanges or through intramolecular exchanges. As shown in Figure 2, a sister-strand crossover between *nif* regions a and b leads to the formation of a large tandem duplication, where the join point carries the *nifD*::Km/Sp allele but lacks the *nifH*::*I-SceI* allele. Alternatively, removal of the *nifH*::*I-SceI* allele by sister-strand exchanges can come about from either double recombination or gene conversion. Both processes generate a non-rearranged pSym, lacking the *nifH*::*I-SceI* allele but maintaining the *nifD*::Km/Sp allele (Figure 2). This kind of product was called apparent gene conversion, since a formal distinction of which process is participating (double recombination or gene conversion) is not possible when the recombining sequences are in direct orientation (Segall and Roth 1994).

Intramolecular exchanges might also be responsible for the generation of selectable Km^r derivatives. As shown in Figure 3 (left part), an intramolecular crossover generates a true gene conversion recombinant. However, this event cannot be distinguished from the non-rearranged class generated by sister-strand exchanges (see above); therefore, all these are also scored as apparent gene convertants. Intramolecular exchange between *nif* regions a and c generates an inversion (Figure 3, center). This recombinant has a Km^r phenotype

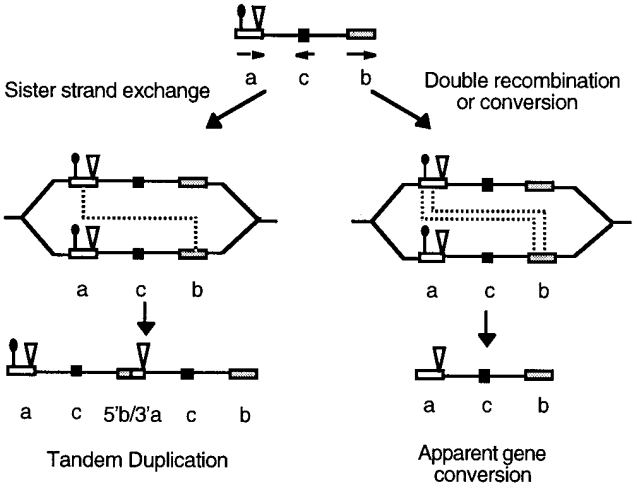

Figure 2.—Formation of Km^r derivatives by sister-strand exchanges. *nif* regions a, b and c are shown as rectangles with white, stippled or black shading, respectively, the *nifH::I-SceI* allele as a stick and ball symbol and the *nifD*::Km/Sp allele as a triangle symbol. (Left part) Recombination between *nif* regions a and b leads to the formation of a large tandem duplication, where the join point (indicated as a rectangle with mixed shading) carries the *nifD*::Km/Sp allele but lacks the *nifH::I-SceI* allele. (Right part) Both double recombination or gene conversion can generate a non-rearranged pSym, lacking the *nifH::I-SceI* allele but maintaining the *nifD*::Km/Sp allele. This kind of product was called apparent gene conversion, since making a formal distinction between which process (double recombination or gene conversion) is participating is not possible when the recombining sequences are in direct orientation. Only the selectable products are shown.

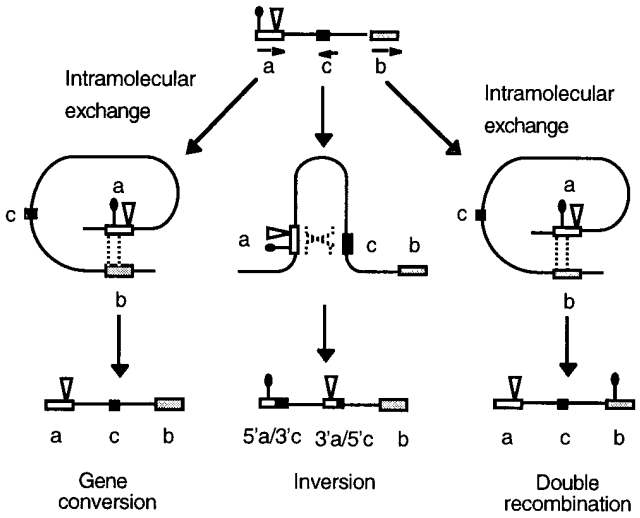

Figure 3.—Intramolecular exchanges might also be responsible for the generation of selectable Km^r derivatives. All symbols are as in Figure 2. (Left part) An intramolecular crossover generates a true gene conversion recombinant. However, this event cannot be distinguished from the non-rearranged class generated by sister-strand exchanges (see Figure 2); therefore, all these are also scored as apparent gene convertants. (Center part) An intramolecular exchange between *nif* regions a and c generates an inversion. This recombinant has a Km^r phenotype due to the relocation of the *nifD*::Km/Sp allele, which is now fused to the *nifH* region c promoter. (Right part) Intramolecular double recombination may also produce a Km^r recombinant, due to a relocation of the *nifH::I-SceI* allele. Only the selectable products are shown.

due to the relocation of the *nifD*::Km/Sp allele, which is now fused to the *nifH* region c promoter. Finally, intramolecular double recombination may also produce a Km^r recombinant, due to a relocation of the *nifH::I-SceI* allele (Figure 3, right).

Correction of a small insertion in the *nif* multigene family is achieved by multiple recombination events: To evaluate the frequency of correction in this multigene family, strain CFNX237, carrying the *nifH::I-SceI* and *nifD*::Km/Sp allelic combination on the pSym, was grown overnight in rich media and plated in media containing a low concentration of kanamycin (3 μg ml^{-1}). Km^r derivatives were found at a high frequency (344×10^{-6}). To identify the molecular events responsible for the generation of Km^r derivatives, 51 colonies were randomly chosen from seven independent selection experiments and purified as single-colony isolates. These derivatives were characterized by a combination of PCR, sizing of the resulting pSym and Southern blot hybridization against specific probes and then assigned to specific classes according to the criteria described in materials and methods.

As shown in Figure 4, derivatives belonging to class I (tandem duplication) were the most abundant, comprising 74% of the observed products. This class may have resulted from a sister-strand exchange (Figure 2, left). Class II derivatives (apparent gene conversions) constitute 14% of the total isolates analyzed. As explained before, these may originate either through interactions involving sister strands (double crossover or gene conversion, Figure 2, right) or from intramolecular exchanges (true gene conversion, Figure 3, left). Inversions (class III, Figure 4) were very scarce in this sample, being represented by a single derivative (*ca.* 2%). This class is readily explained by assuming a reciprocal intramolecular exchange (Figure 3, center; see discussion).

The remaining 10% of the derivatives are divided into three additional classes (IV to VI, Figure 4). These were generated most likely through a combination of apparent gene conversion plus additional recombination events. For instance, class IV is a tandem duplication similar but not identical to class I. Unlike class I, class IV derivatives have additionally lost the *nifH::I-SceI* allele, thus being also a case of apparent gene conversion. Class V is similar to the simple class II derivatives (apparent gene conversion), but it harbors an additional *nifD*::Km/Sp allele in *nif* region b. Class VI is an apparent gene conversion that carries a large deletion of the pSym; deletions similar to this have been observed previously (Romero *et al.* 1991). No derivatives attributable to intramolecular double recombination (Figure 3,

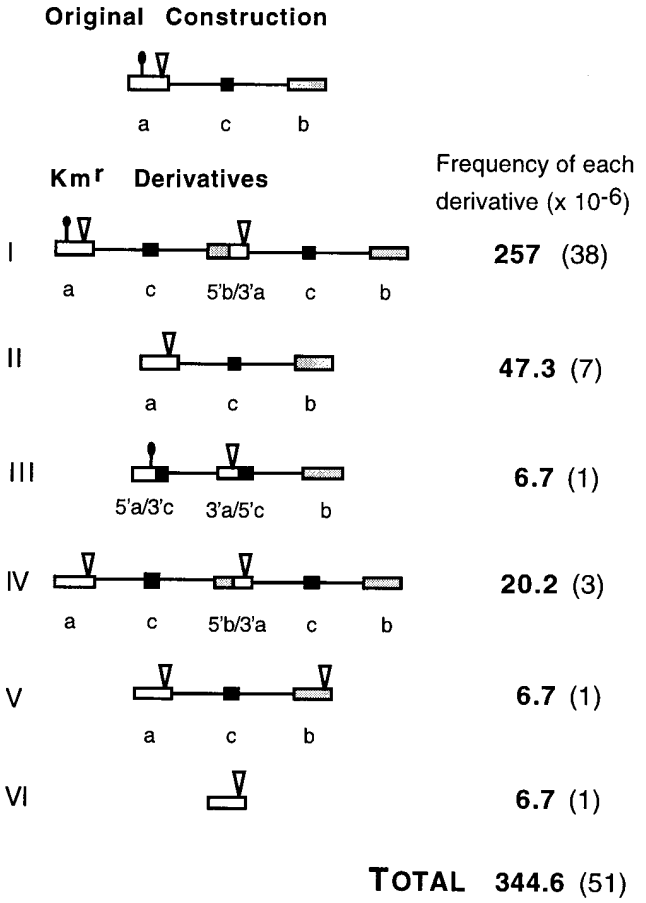

Figure 4.—Molecular events leading to the correction of a small insertion in the *nif* multigene family. Roman numerals indicate specific classes. Numbers in bold denote the frequency of clones in each class, while numbers in parentheses are the number of Km^r clones represented in the corresponding class. All symbols are as in Figure 2.

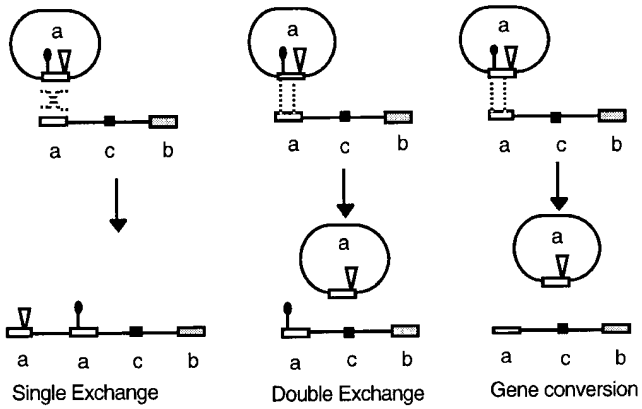

Figure 5.—Formation of Km^r derivatives by intermolecular exchanges. Intermolecular exchanges can form selectable Km^r derivatives either by a double exchange (center part) or by a true gene conversion event (right part). The product shown at the left part (single exchange) is not selectable, because it confers only a low-level resistance to kanamycin.

right) were found, despite the finding of other complex, double-exchange events. Possible reasons for the absence of this class are presented in the discussion.

Thus, these results indicate that: (i) a variety of recombination events are participating in the correction of a small insertion in a member of this multigene family; (ii) tandem duplication is a major contributor for the observed correction, followed significantly by apparent gene conversion; and (iii) multiple recombination events were frequently found.

Role of the *recA* gene in recombinational correction: The participation of the *recA* gene in the processes leading to correction of the *nifH::I-SceI* allele was evaluated by introducing the *recA::*ΩCm allele into strain CFNX237 as described in materials and methods. Introduction of the *recA::*ΩCm allele provokes a 50-fold reduction in the frequency of isolation of Km^r derivatives (from 344×10^{-6} in strain CFNX237 to 6×10^{-6} in strain CFNX242). Characterization of twenty Km^r isolates obtained from strain CFNX242 did not reveal any loss or relocation of the *nifH::I-SceI* allele. Although not characterized further, these derivatives are more likely due to spontaneous mutations that restore transcription of the Km^r gene. Thus, the *recA* gene participates in the formation of tandem duplications and apparent gene conversions.

Recombinational correction in the *nif* multigene family occurs frequently by intermolecular exchanges: A particular limitation of the system employed so far is that correction of the *nifH::I-SceI* allele might be generated either by intermolecular exchanges (sister-strand events), by intramolecular recombination or through a combination of both processes. However, a clear distinction of the role of intermolecular exchanges can be achieved through the incorporation of the allele whose correction is to be scored into a separate plasmid. As shown in Figure 5, correction in this system occurs only by intermolecular exchanges. This process has three separate outcomes. Single exchanges provoke the incorporation of the small plasmid into the pSym (Figure 5, left). The other two alternatives (double exchange or gene conversion) maintain the small plasmid and the pSym as separate entities, but the small plasmid now lacks the *nifH::I-SceI* allele (Figure 5, center and right). Although the expected products for double exchange or gene conversion events are different, it is not feasible to achieve a distinction between these alternatives, due to the possibility of cosegregation between the small plasmid (generated by a double exchange) and a second, uninvolved copy of the pSym.

To evaluate the role of intermolecular exchanges in the observed recombinational correction, small, broad host-range plasmids containing either the *nifD::*Km/Sp allele (pCRS6, Table 1) or the *nifH::I-SceI* and *nifD::*Km/Sp allelic combination (pCRS7, Table 1) were constructed as described in materials and methods. Conjugative transfer of each plasmid into a wild-type *R. etli*

strain (CE3) generated strains CFNX238 and CFNX239 (Table 1). Strain CFNX238 (harboring plasmid pCRS6) showed a high-level resistance to kanamycin (70 μg ml^{-1}), while strain CFNX239 (carrying pCRS7) remained Km^s. Km^r derivatives from strain CFNX239 were obtained at a frequency of 28×10^{-6}. A total of 21 separate isolates (obtained from four independent selection experiments) were classified under the criteria described in materials and methods. All these derivatives showed a structure compatible with apparent gene conversion (Figure 5, right). The absence of the single-exchange class was expected, due to the low resistance to kanamycin observed when the *nifH::I-SceI* and *nifD::*Km/Sp allelic combination was integrated on the pSym. Similar to the previous system, the double-exchange class was absent from this sample (see discussion).

The frequency of the apparent gene conversion class in this system (28×10^{-6}) is similar to the one estimated before (48×10^{-6}). These results suggest that this class is frequently generated through intermolecular exchanges.

This system allows us to determine if the formation of apparent gene convertants depends on the presence of additional copies of the *nif* region. To that end, plasmid pCRS7 was introduced into strain CFNX89, generating strain CFNX241; since this strain is devoid of the pSym, additional copies of the *nif* region are absent. The frequency of Km^r derivatives is reduced threefold (to 7.1×10^{-6}); these derivatives did not show the correction of the *nifH::I-SceI* allele. Thus, these results, coupled to the *recA* dependency data, indicate that apparent gene conversions are formed through a recombinational process, and not due to a *recA*-dependent, excision repair process.

Correction in the *nif* multigene family leads to a precise restoration in nucleotide sequence: To be useful as a process leading to concerted variation between the *nif* regions, apparent gene conversion must restore the *nifH::I-SceI* allele to an otherwise wild-type sequence. To ascertain if this was the case, three independent class II, Km^r derivatives and one class VI derivative (Figure 4) were subjected to PCR amplifications employing oligonucleotides o1 and o4 (Figure 1) as primers. This particular primer pair yielded products of the region encompassing the site where the *nifH::I-SceI* allele was located. These PCR products were sequenced, using oligonucleotide o1 as a primer. In every case, nucleotide sequence was fully restored to wild type. These data were extended by isolating pCRS7 from four independent Km^r derivatives obtained as explained in the previous section; these plasmids were then transformed into *E. coli*. Nucleotide sequence from these derivatives also yielded a region identical to the wild-type sequence (data not shown). Thus, apparent gene conversions lead to a precise restoration of the wild-type nucleotide sequence.

DISCUSSION

In this study we show that multiple recombination events, including tandem duplication, apparent gene conversion and inversion, lead to the correction of a small insertion in the *nif* multigene family. Tandem duplications appeared at a frequency of 2.8×10^{-4}, and were formed through recombination between *nif* region a and *nif* region b. The abundance of this class (75% of the observed events) was expected, since any crossover in the 1-kb region separating the *nifH::I-SceI* and *nifD::*Km/Sp alleles must produce a Km^r derivative. Thus, the frequency at which this class is generated reflects the frequency of recombination per kb of available homology in this organism. This estimate agrees well with previous evaluations in an interval of equivalent size (Romero *et al.* 1995).

Apparent gene conversion events in the *nif* multigene family were readily detected. The recombinational origin of these events is supported by (i) the strict requirement for a functional *recA* gene, (ii) the dependency on additional copies of the *nif* region for their formation, and (iii) a precise restoration of the wild-type nucleotide sequence. Apparent gene conversions comprised 24% of the observed events, appearing at a cumulative frequency of 8×10^{-5}. This estimate is remarkably similar to the observed frequencies of apparent gene conversion in both *E. coli* (Harvey and Hill 1990) and *S. typhimurium* (Segall and Roth 1994). Previous estimates obtained in *R. meliloti*, based on data from phage crosses, suggested that the efficiency of apparent gene conversion was lower in this organism than in *E. coli* (Orosz *et al.* 1980). Our results clearly show that this is not the case, at least for recombinational interactions among members of a multigene family.

It is important to remark that this frequency refers to interactions with a particular outcome, namely, the correction of a specific 28-bp insertion. When all the possible corrections alongside the 5-kb *nif* regions are considered, the fraction of cells in a population with some change caused by apparent gene conversion in the *nif* regions should become substantial. Our results show that transfer of genetic information through apparent gene conversion occurs at a frequency that exceeds the frequency of spontaneous mutation. Therefore, this mechanism must contribute significantly to maintain the genetic homogeneity observed among members of the *nif* multigene family in Rhizobium, resulting in concerted evolution. We suggest that apparent gene conversion should also be relevant to explain other instances of maintenance of sequence identity in Rhizobium, as seen in a variety of multigene families.

Besides demonstrating the occurrence of apparent gene conversion, our work also allows us to draw some useful inferences about the recombination process in Rhizobium. Compared to tandem duplication and ap-

parent gene conversion, inversions are minor contributors to correction of the effects of a small insertion. Only a single inversion, formed by recombination between *nif* region a and *nif* region c was isolated, at a frequency of 6×10^{-6}. Although the ocurrence of an inversion allows us to classify this interval as permissive (according to the terminology of Segall and Roth 1994), the scarcity of this class is intriguing. Possible deleterious effects of this rearrangement on strain viability can be ruled out, because the strain harboring an inversion grows as well as the parental strain (data not shown). The infrequent occurrence of inversions cannot be explained by restriction in the size of a homologous region available for recombination. The interval used to generate an inversion is similar in location and size (750 bp) to the one used to generate a tandem duplication (1 kb). Despite this similarity, inversions occur 100-fold less frequently than tandem duplications.

Rarity of the inversion class can be explained by assuming, as has been proposed previously, that recombination in bacteria frequently entails the use of the so-called half-crossing over (Mahan and Roth 1989; Kobayashi 1992; Segall and Roth 1994; Roth *et al.* 1996; Yamamoto *et al.* 1996). Following this proposal, a half-crossover between two recombining sequences would generate a recombinant molecule and two broken DNA ends. Tandem duplications and apparent gene conversions can be readily generated by this mechanism; in contrast, formation of an inversion would require two coincidental half-crossovers to produce a viable recombinant, explaining the low frequency of this class. The occurrence of half-crossovers would also explain two additional aspects of this work, namely the preference of intermolecular exchanges for generation of apparent gene convertants and the absence of derivatives attributable to intramolecular double recombination (Figure 3, right). These features might be a consequence of the fact that a single intramolecular half-crossover generates a linear molecule, thus precluding the isolation of a viable recombinant.

Considering the arguments given above, the occurrence of classes attributable to complex, double-exchange events (classes IV to VI, Figure 4) might seem paradoxical. However, we believe that these classes are formed by successive, rather than coincidental, half-crossover events. Common to these classes is that they appear to be composites of two different recombination events, where the product of each event is selectable. For instance, class IV was apparently generated by two selectable events, one a tandem duplication and the other an apparent gene conversion. It is conceivable that recombinants belonging to this class may have formed as successive recombination events during colony development, in a situation akin to phage recombination or recombination among small, high-copy number plasmids (Yamamoto *et al.* 1988).

We are indebted to Susana Brom, Jaime Martínez-Salazar and Rafael Palacios for critical reviewing of the manuscript, to José Espíritu and César Hernández for help in preparing the manuscript, to Patricia Bustos for help with nucleotide sequencing, and to Javier Rivera and María de la Paz Salas for skillful technical assistance. Partial financial support for this research was provided by grants No. 4321-N9406 from the Consejo Nacional de Ciencia y Tecnología (México) and No. 030355 from the Programa de Apoyo a las Divisiones de Estudios de Posgrado (Universidad Nacional Autónoma de México). C. R. was a recipient of a scholarship from the Consejo Nacional de Ciencia y Tecnología (México).

LITERATURE CITED

Abdulkarim, F., and D. Hughes, 1996 Homologous recombination between the *tuf* genes of *Salmonella typhimurium*. J. Mol. Biol. **260:** 506–522.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman *et al.*, 1987 *Current Protocols in Molecular Biology*. Massachusetts General Hospital, Harvard Medical School. John Wiley & Sons, Inc., New York.

Bachellier, S., E. Gilson, M. Hofnung and C. W. Hill, 1996 Repeated sequences, pp. 2012–2040 in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, Ed. 2, edited by F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, Jr. *et al.* American Society for Microbiology Press, Washington, DC.

Badenoch-Jones, J., T. A. Holton, C. Morrison, K. F. Scott and J. Shine, 1989 Structural and functional analysis of nitrogenase genes from the broad-host-range *Rhizobium* strain ANU240. Gene **77:** 141–153.

Bergman, K., E. Nulty and L. Su, 1991 Mutations in the two flagellin genes of *Rhizobium meliloti*. J. Bacteriol. **173:** 3716–3723.

Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland *et al.*, 1997 The complete genome sequence of *Escherichia coli* K-12. Science **277:** 1453–1474.

Boyer, H. W., and D. Roulland-Dussoix, 1969 A complementation analysis of restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41:** 459–472.

Brom, S., A. García de los Santos, T. Stepkowski, M. Flores, G. Dávila *et al.*, 1992 Different plasmids of *Rhizobium leguminosarum* bv. phaseoli are required for optimal symbiotic performance. J. Bacteriol. **174:** 5183–5189.

Buchanan-Wollaston, V., M. C. Cannon, J. L. Beynon and F. C. Cannon, 1981 Role of the *nifA* gene product in the regulation of *nif* expression in *Klebsiella pneumoniae*. Nature **294:** 776–778.

David, M., O. Domergue, P. Pognonec and D. Kahn, 1987 Transcription patterns of *Rhizobium meliloti* symbiotic plasmid pSym: identification of *nifA*-independent *fix* genes. J. Bacteriol. **169:** 2239–2244.

Deonier, R. C., 1996 Native insertion sequence elements: locations, distributions and sequence relationships, pp. 2000–2011 in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, Ed. 2, edited by F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, Jr. *et al.* American Society for Microbiology Press, Washington, DC.

Dover, G. A., 1993 Evolution of genetic redundancy for advanced players. Curr. Opin. Genet. Dev. **3:** 902–910.

Eckhardt, T., 1978 A rapid method for the identification of plasmid deoxyribonucleic acid in bacteria. Plasmid. **1:** 584–588.

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

Figurski, D. H., and D. R. Helinski, 1979 Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. Proc. Natl. Acad. Sci. USA **76:** 1648–1652.

Fischer, H.-M., M. Babst, T. Kaspar, G. Acuña, F. Arigoni *et al.*, 1993 One member of a *groESL*-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. EMBO J. **12:** 2901–2912.

Flores, M., V. González, S. Brom, E. Martínez, D. Piñero *et al.*, 1987 Reiterated DNA sequences in *Rhizobium* and *Agrobacterium* spp. J. Bacteriol. **169:** 5782–5788.

Freiberg, C., R. Fellay, A . Bairoch, W. J. Broughton, A. Rosenthal *et al.*, 1997 Molecular basis of symbiosis between *Rhizobium* and legumes. Nature **387:** 394–401.

Gangloff, S., H. Zou and R. Rothstein, 1996 Gene conversion plays the major role in controlling the stability of large tandem repeats in yeast. EMBO J. **15:** 1715–1725.

Girard, M. L., M. Flores, S. Brom, D. Romero, R. Palacios *et al.*, 1991 Structural complexity of the symbiotic plasmid of *Rhizobium leguminosarum* bv. phaseoli. J. Bacteriol. **173:** 2411–2419.

Harvey, S., and C. W. Hill, 1990 Exchange of spacer regions between rRNA operons in *Escherichia coli.* Genetics **125:** 683–690.

Hernández-Lucas, I., M. A. Pardo, L. Segovia, J. Miranda and E. Martínez-Romero, 1995 *Rhizobium tropici* chromosomal citrate synthase gene. Appl. Environ. Microbiol. **61:** 3992–3997.

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

John, B., and G. L. G. Miklos, 1988 *The Eukaryote Genome in Development and Evolution.* Allen & Unwin, London.

Jones, J. D. G., and N. Gutterson, 1987 An efficient mobilizable cosmid vector, pRK7813, and its use in a rapid method for marker exchange in *Pseudomonas fluorescens* HV37a. Gene **61:** 299–306.

Klein, H. L., and T. D. Petes, 1981 Intrachromosomal gene conversion in yeast. Nature **289:** 144–148.

Kobayashi, I., 1992 Mechanisms for gene conversion and homologous recombination: the double-strand break repair model and the successive half crossing-over model, pp. 81–133 in *Advances in Biophysics*, edited by M. Kotani. Japan Scientific Societies Press and Elsevier, Tokyo and Limerick.

Kündig, C., H. Hennecke and M. Göttfert, 1993 Correlated physical and genetic map of the *Bradyrhizobium japonicum* 110 genome. J. Bacteriol. **175:** 613–622.

Mahan, M. J., and J. R. Roth, 1989 Role of *recBC* function in formation of chromosomal rearrangements: a two-step model for recombination. Genetics **121:** 433–443.

Margolin, W., and S. Long, 1994 *Rhizobium meliloti* contains a novel second homolog of the cell division gene *ftsZ.* J. Bacteriol. **176:** 2033–2043.

Marsh, J. L., M. Erfle and E. J. Wykes, 1984 The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene **32:** 481–485.

Martínez-Salazar, J. M., D. Romero, M. L. Girard and G. Dávila, 1991 Molecular cloning and characterization of the *recA* gene of *Rhizobium phaseoli* and construction of *recA* mutants. J. Bacteriol. **173:** 3035–3040.

Miller, J. H., 1972 *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Morett, E., S. Moreno and G. Espín, 1988 Transcription analysis of the three *nifH* genes of *Rhizobium phaseoli* with gene fusions. Mol. Gen. Genet. **213:** 499–504.

Noel, K. D., F. Sánchez, L. Fernández, J. Leemans and M. A. Cevallos, 1984 *Rhizobium phaseoli* symbiotic mutants with transposon Tn*5* insertions. J. Bacteriol. **158:** 148–155.

Norel, F., and C. Elmerich, 1987 Nucleotide sequence and functional analysis of the two *nifH* copies of *Rhizobium* ORS571. J. Gen. Microbiol. **133:** 1563–1576.

Ohno, S., 1970 *Evolution by Gene Duplication.* Springer-Verlag, Berlin.

Okazaki, N., S. Matsuo, K. Saito, A. Tominaga and M. Enomoto, 1993 Conversion of the *Salmonella* phase 1 flagellin gene *fliC* to the phase 2 gene *fljB* on the *Escherichia coli* K-12 chromosome. J. Bacteriol. **175:** 758–766.

Orosz, L., A. Páy and G. Dallmann, 1980 Heterozygosis of phage *16-3* of *Rhizobium meliloti*: moderate level of mismatch repair or gene conversion. Mol. Gen. Genet. **179:** 163–167.

Pardo, M. A., J. Lagúnez, J. Miranda and E. Martínez, 1994 Nodulating ability of *Rhizobium tropici* is conditioned by a plasmid-encoded citrate synthase. Mol. Microbiol. **11:** 315–321.

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

Prentki, P., and H. M. Krisch, 1984 In vitro insertional mutagenesis with a selectable DNA fragment. Gene **29:** 303–313.

Quinto, C., H. de la Vega, M. Flores, J. Leemans, M. A. Cevallos *et al.*, 1985 Nitrogenase reductase: a functional multigene family in *Rhizobium phaseoli.* Proc. Natl. Acad. Sci. USA **82:** 1170–1174.

Romero, D., S. Brom, J. Martínez-Salazar, M. L. Girard, R. Palacios *et al.*, 1991 Amplification and deletion of a *nod-nif* region in the symbiotic plasmid of *Rhizobium phaseoli.* J. Bacteriol. **173:** 2435–2441.

Romero, D., J. Martínez-Salazar, L. Girard, S. Brom, G. Dávila *et al.*, 1995 Discrete amplifiable regions (amplicons) in the symbiotic plasmid of *Rhizobium etli* CFN42. J. Bacteriol. **177:** 973–980.

Romero, D., G. Dávila and R. Palacios, 1997 The dynamic genome of *Rhizobium*, pp. 153–161 in *Bacterial Genomes: Physical Structure and Analysis*, edited by F. J. de Bruijn, J. R. Lupski and G. Weinstock. Chapman & Hall, New York.

Roth, J. R., N. Benson, T. Galitski, K. Haack, J. G. Lawrence *et al.*, 1996 Rearrangements of the bacterial chromosome: formation and applications, pp. 2256–2276 in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, Ed. 2, edited by F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, Jr. *et al.* American Society for Microbiology Press, Washington, DC.

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

Schlüter, A., T. Patschkowski, J. Quandt, L. B. Selinger, S. Weidner *et al.*, 1997 Functional and regulatory analysis of the two copies of the *fixNOQP* operon of *Rhizobium leguminosarum* strain VF39. Mol. Plant Microbe Interact. **10:** 605–616.

Schultze, M., E. Kondorosi, P. Ratet, M. Buiré and A. Kondorosi, 1994 Cell and molecular biology of *Rhizobium*-plant interactions. Int. Rev. Cytol. **156:** 1–75.

Segall, A. M., and J. R. Roth, 1994 Approaches to half-tetrad analysis in bacteria: recombination between repeated, inverse-order chromosomal sequences. Genetics **136:** 27–39.

Segovia, L., J. P. W. Young and E. Martínez-Romero, 1993 Reclassification of american *Rhizobium leguminosarum* biovar phaseoli type I strains as *Rhizobium etli sp. nov.* Int. J. Sist. Bacteriol. **43:** 374–377.

Simon, R., 1984 High frequency mobilization of gram-negative bacterial replicons by the *in vitro* constructed Tn*5*-Mob transposon. Mol. Gen. Genet. **196:** 413–420.

Simon, R., U. Priefer and A. Pühler, 1983 A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram-negative bacteria. Bio/Technology **1:** 784–791.

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

Valderrama, B., A. Dávalos, L. Girard, E. Morett and J. Mora, 1996 Regulatory proteins and *cis*-acting elements involved in the transcriptional control of *Rhizobium etli* reiterated *nifH* genes. J. Bacteriol. **178:** 3119–3126.

Williams, S. M., and C. Strobeck, 1985 Sister chromatid exchange and the evolution of rDNA spacer length. J. Theor. Biol. **116:** 625–636.

Yamamoto, K., H. Yoshikura, N. Takahashi and I. Kobayashi, 1988 Apparent gene conversion in an *Escherichia coli* rec^+ strain is explained by multiple rounds of reciprocal crossing over. Mol. Gen. Genet. **212:** 393–404.

Yamamoto, K., N. Takahashi, Y. Fujitani, H. Yoshikura and I. Kobayashi, 1996 Orientation dependence in homologous recombination. Genetics **143:** 27–36.

Communicating editor: P. L. Foster