Mol Gen Genet (1988) 212:6–10

MGG

The chloroplast genomes of conifers lack one of the rRNA-encoding inverted repeats

Jonas Lidholm[1], Alfred E. Szmidt[2], Jan-Erik Hällgren[2], and Petter Gustafsson[1]

[1] Plant Molecular Biology Group, Institute of Cell and Molecular Biology, University of Umeå, S-901 87 Umeå, Sweden
[2] Department of Forest Genetics and Plant Physiology, Faculty of Forestry, Swedish University of Agricultural Sciences, S-901 83 Umeå, Sweden

Summary. Chloroplast DNA from species of five different conifer genera was extracted and studied by Southern blot analysis. For all these species, hybridization with heterologous probes specific for 16 S and 23 S rDNA detected only one chloroplast DNA fragment per enzyme digest. This observation suggests that the 16 S and 23 S rRNA genes are not duplicated in these genomes. The unique 16 S rDNA-containing *Bam*HI fragment from *Pinus contorta* Dougl. was cloned and restriction mapped. Apart from the 16 S rRNA gene, this fragment also contained the *psbC* and *psbD* genes. It is concluded that the chloroplast genomes of a wide taxonomic range of conifers lack one of the inverted repeat elements and that a dislocation of the *psbDC* gene cluster has occurred in *P. contorta*.

Key words: Chloroplast genome – Inverted repeat – Conifers – Gymnosperms

Introduction

Chloroplasts possess their own genetic system which encodes and expresses a substantial part of the protein components involved in the photosynthetic process. The chloroplast (cp) genomes from a number of angiosperms, mosses and algae have been extensively characterized (reviewed by Palmer 1985b). The homogeneity among these genomes with respect to gene content, structural organization and size is remarkable. The cp genome consists of a circular molecule which averages about 150 kb in size and exists as multiple copies per plastid. The gross structure of the cp genome normally comprises two rDNA-containing inverted repeats (IR) which divide the circular molecule into two single-copy regions. The only deviations from this organization found to date among higher plants are some members of the *Fabaceae* family, which lack one of the repeats (reviewed by Palmer 1985a). Among all the cp genomes studied so far, the representation of gymnosperms is very poor. To our knowledge, no structural study of conifer cp genomes has yet been published, although a few studies on preparation and restriction patterns of cpDNA from pines (Kondo et al. 1986; Szmidt et al. 1986; White 1986) were recently reported. It has also been shown that cpDNA is inherited paternally in conifers (Neale et al. 1986; Szmidt et al. 1987; Wagner et al. 1987).

Offprint requests to: J. Lidholm

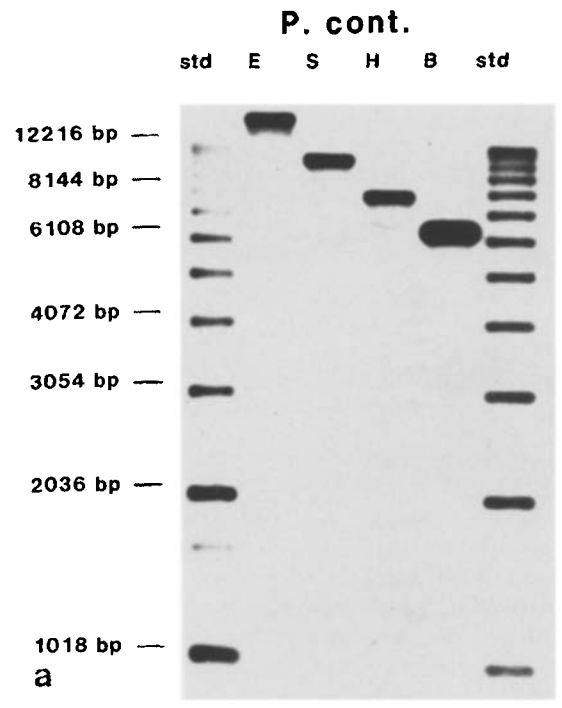
In this paper we describe the results of heterologous hybridization studies of cpDNA from representatives of five different genera of conifers. We show that all these cp genomes contain the same deviation from the 'consensus' structure of higher plants, as do some of the *Fabaceae* plants, i.e. absence of one of the normally repeated rRNA-encoding segments.

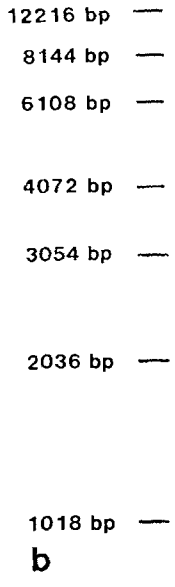
Materials and methods

Plant material, bacteria, cloning vector and probes. Needles were collected from seedlings and individual trees of *Pinus contorta* Dougl., *Picea sitchensis* (Bong.) Carr., *Larix sibirica* Ledeb., *Abies balsamea* (L.) Mill. and *Thuja standishii* (Gord.) Carr. The plasmid used for the library was pUC9 (Yanisch-Perron et al. 1985). The *Escherichia coli* strain used was DH5α, a derivative of DH1 (Hanahan 1983). Probes specific for 16 S and 23 S rDNA were derived from the 6.6 kb *Pst*I insert of the clone pAN4, which contains rDNA from *Anacystis nidulans* (Tomioka and Sugiura 1983, a gift from Dr. M. Sugiura). The 1.34 kb 16 S probe ranged from the *Sac*I site at nucleotide 198 of the 16 S rRNA gene to the *Hin*dIII site 50 nucleotides 3′ of the coding sequence. The 0.87 kb 23 S probe was an internal *Hin*dIII fragment of the 23 S rRNA gene, ranging from position 730 to 1600. The *psbC* and *psbD* probes were the spinach clones pPSII44 and pPSII32/2, respectively (Alt et al. 1984, gifts from Dr. R.G. Herrmann).

DNA techniques. The method for extraction of cpDNA from pines has been described elsewhere (Szmidt et al. 1986). DNA was digested with restriction enzymes according to standard procedures. Samples were loaded on 0.7% agarose gels and electrophoresis was carried out in TAE-buffer (Maniatis et al. 1982). Blot transfer to nylon membrane was made according to Southern (1975). The DNA size marker used was the BRL 1 kb ladder which comprises a ladder of 1018 bp increments, ranging from 1018 to 12216 bp. Radioactive probes were prepared using the random priming procedure (Feinberg and Vogelstein 1983). Heterologous filter hybridizations were performed according to Maniatis et al. (1982) to obtain 75% stringency level. To remove probe prior to rehybridization, filters were washed in 0.1 × SSC, 0.1% SDS at 95° C for 10 min.

A cpDNA library was made by ligating 1 μg of *Bam*HI-digested cpDNA from *P. contorta* to pUC9 and transforming DH5α cells by the RbCl method of Hanahan (1985).





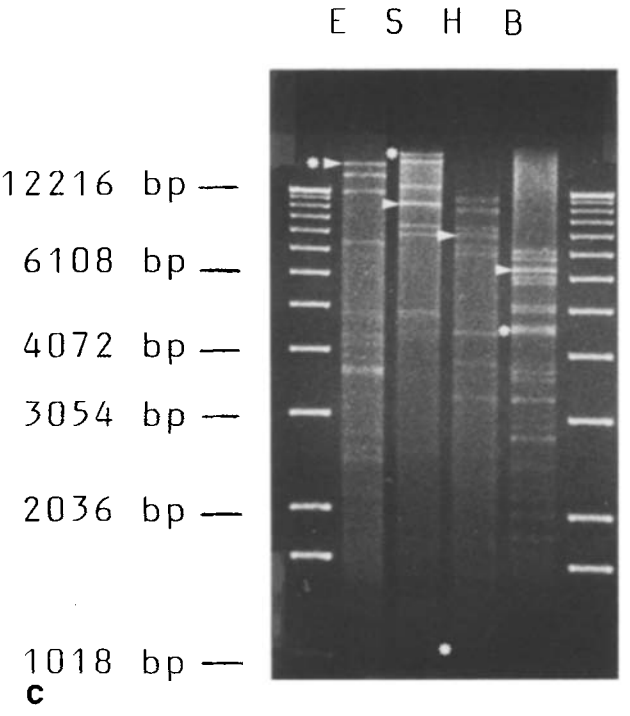


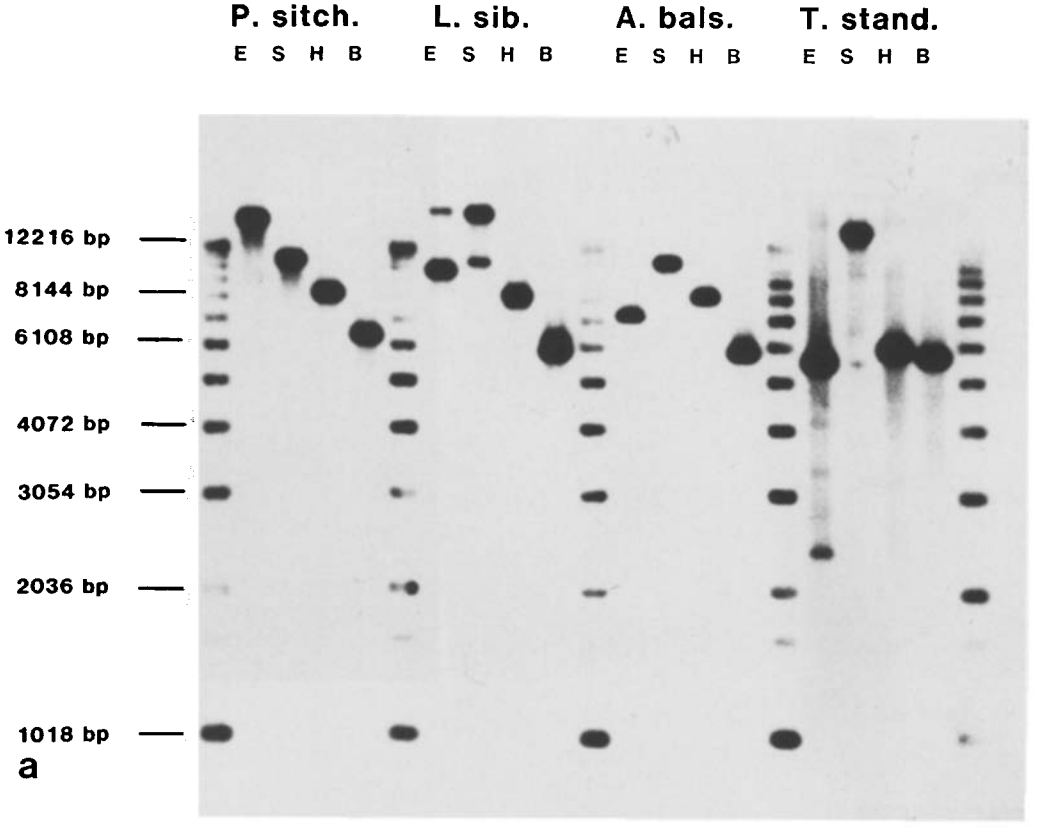
Fig. 1a–c. Southern blot hybridizations of *Pinus contorta* chloroplast DNA using **a** the 16 S and **b** the 23 S rDNA probes; **c** agarose gel. The digests are *Eco*RI (E), *Sac*I (S), *Hin*dIII (H) and *Bam*HI (B). The lanes marked 'std' contain the molecular weight marker. In **c**, *arrowheads* and *asterisks* indicate bands hybridizing to the 16 S and the 23 S rDNA probes, respectively

Colony hybridization (Maniatis et al. 1982), using the 16 S rDNA probe, was performed to identify recombinant clones containing the 16 S rRNA gene. The isolated 16 S-positive clone, pPC28, was restriction mapped by conventional methods. Plasmids were prepared according to the alkaline lysis method described in Maniatis et al. (1982).

Results and discussion

The genes encoding 16 S and 23 S rRNA are normally located close to the borders of the IR elements. One would therefore expect them generally to occur on fragments covering the junctions of the repeats and the adjacent single-copy region, i.e. at least two bands per digest would be seen in filter hybridizations with rDNA probes. When hybridized to a Southern blot of *P. contorta* cpDNA, each of the probes specific for 16 S and 23 S rDNA detected only one fragment per digest (Fig. 1a, b). Three of these fragments, the *Sac*I and *Bam*HI fragments detected by the 16 S probe and the *Bam*HI fragment detected by the 23 S probe, correspond to bands of increased intensity on the agarose gel, i.e. bands consisting of two or more co-migrating fragments. The other hybridizing fragments can be identified as one-molar bands (Fig. 1c).

We interpret the observed hybridization patterns as revealing the presence of only one set of rRNA genes on this cp genome. The alternative explanation would be that there is indeed an IR organization and that the double



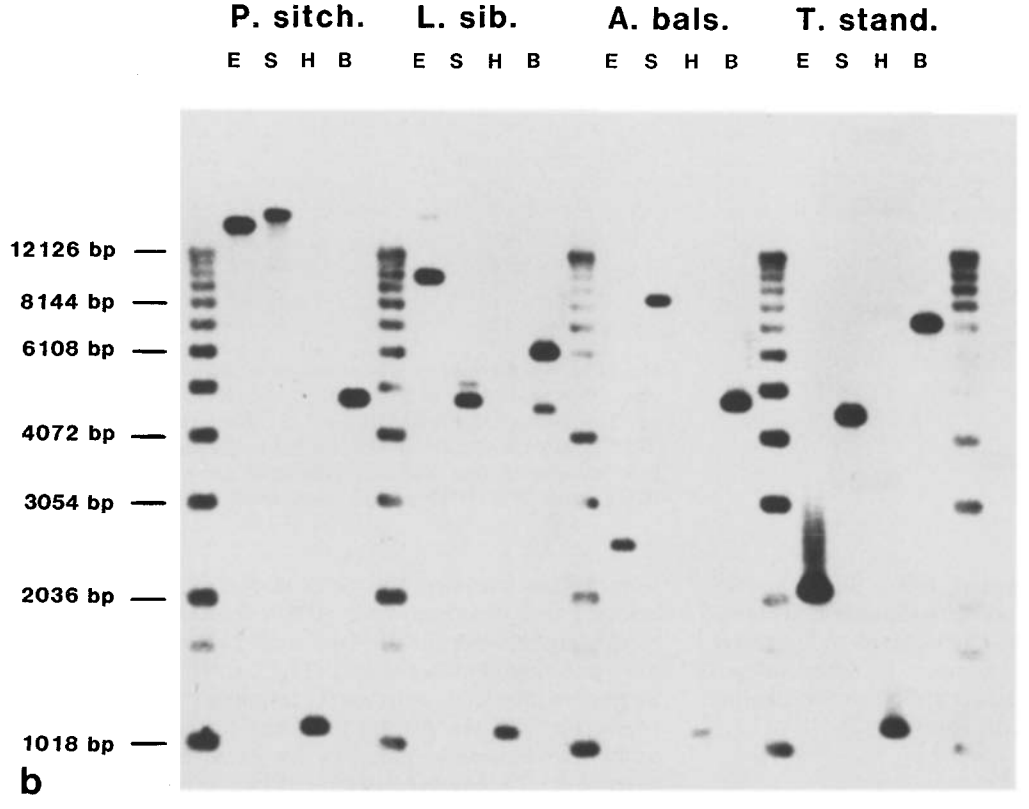


Fig. 2a, b. Southern blot hybridizations of chloroplast DNA from four different genera of conifers, using **a** the 16 S and **b** the 23 S rDNA probes. The restriction enzymes used are *Eco*RI (E), *Sac*I (S), *Hin*dIII (H) and *Bam*HI (B)

bands which hybridize to the rDNA probes are co-migrating fragments from either within the IR or from the junction of the IR/single-copy region, accidentally being of the same size. This requires that each of the large, one-molar *Eco*RI and *Sac*I fragments spans over one entire single-copy region from corresponding positions in the two repeats. This explanation cannot be true, however, since the *Sac*I fragment is considerably larger than the *Eco*RI fragment, which har-

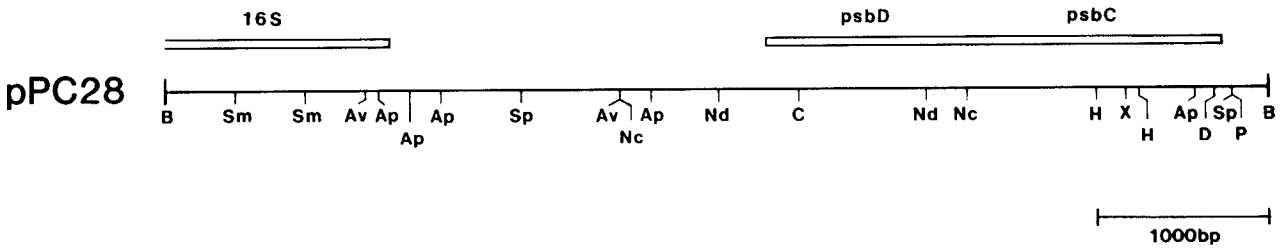


Fig. 3. Restriction map of the cloned *Pinus contorta* fragment pPC28. Recognition sites are indicated for *Apa*I (Ap), *Ava*I (Av), *Bam*HI (B), *Cla*I (C), *Dra*I (D), *Hin*dIII (H), *Nco*I (Nc), *Nde*I (Nd), *Pst*I (P), *Sma*I (Sm), *Sph*I (Sp) and *Xba*I (X)

bours both the 16 S and the 23 S rRNA genes, and yet it lacks the 16 S rRNA gene (Fig. 1a, b).

To find out whether our unexpected result was specific for *P. contorta* or if it might be general for conifers, we extracted cpDNA from representatives of four other conifer genera: *Picea, Larix, Abies* and *Thuja* and performed the same Southern blot analyses on this material. Again, the 16 S (Fig. 2a) and the 23 S (Fig. 2b) rDNA probes detected only one restriction fragment each per digest. In some of the digests of *L. sibirica* and *T. standishii* cpDNA (Fig. 2a, b), one additional hybridizing band can be seen. Since these bands are much weaker than the respective major bands, we assume that they are due either to restriction sites occurring within the probe-homologous sequences or to the digestions of these samples not being absolutely complete.

The only group of higher plants that is so far known to lack one of the normally duplicated rRNA gene clusters is one section of the subfamily *Papilionideae* of the family *Fabaceae* (reviewed by Palmer 1985a). This group contains genera such as *Pisum, Vicia, Lathyrus, Trifolium* and *Melilotus*. Since the study presented here comprises representatives from a wide taxonomical range of conifers, we find it reasonable to believe that lack of one of the IR elements is truly general among these plants. This would make conifers by far the largest group of plants to possess this cp genome organization. It may also encourage some discussion concerning the evolutionary direction with respect to the IR structure. However, we prefer to consider both the conifer and the *Papilionideae* cp genome organizations as the results of rearrangements.

It is of considerable interest to note that *Ginkgo biloba,* which is an ancient gymnosperm, has retained the IR structure (Palmer and Stein 1987). This cp genome is also significantly larger (158 kb) than that of all conifer species studied to date (approximately 123 kb; Kondo et al. 1986; Szmidt et al. 1986). As to when the event occurred which eliminated one of the IR segments in conifer cp genomes, one must consider the time for the divergence of the *Ginkgo* and the conifer lineages as well as allow for the phylogenetic age of the *Pinateae* class; more than 200 million years (Mirov 1967).

Concerning the function of the IR structure, Palmer and Thompson (1982) proposed that absence of one of the repeat elements confers phylogenetic instability on the cp genome. Although data from different conifer species, either to verify or contradict this hypothesis, are lacking, we believe that conifers would be a suitable group of plants to use for this purpose.

From other hybridization studies of the *P. contorta* cp genome (not shown) it appeared that the 16 S rRNA, the *psbC* and the *psbD* genes were located closely together. To obtain an accurate description of this region of the genome, we made a cp DNA library and isolated the 6.5 kb *Bam*HI fragment that hybridized to the probes for these genes. The restriction site map for this fragment, designated pPC28, was established and the three genes were positioned (Fig. 3). Comparison of the cp genomes of *Pisum sativum* (Rasmussen et al. 1984) and *P. contorta* reveals a further common deviation from the 'higher plant consensus': the *psbD* and *psbC* genes are shifted from a position in the middle of the large single-copy region to a location in close proximity to the rRNA genes. In *Pisum,* however, the *psbD* and *psbC* genes neighbour the 23 S rRNA gene whereas in *P. contorta* they are located on the 16 S side of the rRNA gene cluster.

At present we do not know whether the entire content of one repeat segment is lost in the conifers or if the loss is confined only to the rRNA genes. By continued hybridization analysis we intend to investigate this. It will also be interesting to study the cp genome of *Taxaceae* species, which are believed to have diverged from the *Pinateae* lineage at an early stage (Dallimore et al. 1966).

Based on the finding presented in this report, we suggest that conifers should be recognized as a distinct lineage among vascular plants with respect to chloroplast genome evolution.

Acknowledgements. We are grateful to M. Burström, T. Norberg and A. Sigurgeirsson for skilful technical assistance; to Dr. R.G. Herrmann and Dr. M. Sugiura for providing us with the probes used in this study; and to the staff at Arboretum Hoersholm, Denmark for providing plant material. This work was supported by grants from the Swedish Council for Forestry and Agricultural Research and from Bo Rydins Foundation for Scientific Research.

References

Alt J, Morris J, Westhoff P, Herrmann RG (1984) Nucleotide sequence of the clustered genes for the 44 kd chlorophyll *a* apoprotein and the "32 kd"-like protein of the photosystem II reaction center in the spinach plastid chromosome. Curr Genet 8:597–606

Dallimore W, Jackson AB, Harrison SG (1966) A handbook of *Coniferae* and *Ginkgoaceae*. Edward Arnold, London

Feinberg AP, Vogelstein BA (1983) A technique for radiolabelling DNA endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Hanahan D (1983) Studies on the transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580

Hanahan D (1985) Techniques for transformation of *E. coli*. In: Glover DM (ed) DNA cloning, vol 1, a practical approach. IRL Press, Oxford, pp 109–135

Kondo T, Ishibashi T, Shibata M, Hirai A (1986) Isolation of chloroplast DNA from *Pinus*. Plant Cell Physiol 27:741–744

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbour Laboratory Press, New York

Mirov NT (1967) The genus *Pinus*. The Ronald Press, New York

Neale DB, Wheeler NC, Allard RW (1986) Paternal inheritance of chloroplast DNA in Douglas-fir. Can J Forest Res 16:1152–1154

Palmer JD (1985a) Evolution of chloroplast and mitochondrial DNA in plants and algae. In: MacIntyre RJ (ed) Monographs in evolutionary biology: Molecular evolutionary genetics. Plenum, New York, pp 131–240

Palmer JD (1985b) Comparative organization of chloroplast genomes. Annu Rev Genet 19:325–354

Palmer JD, Stein DB (1987) Conservation of chloroplast genome structure among vascular plants. Curr Genet 10:823–833

Palmer JD, Thompson WF (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell 29:537–550

Rasmussen OF, Bookjans G, Stummann BM, Henningsen KW (1984) Localization and nucleotide sequence of the gene for the membrane polypeptide D2 from pea chloroplast DNA. Plant Mol Biol 3:191–199

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Szmidt AE, Lidholm J, Hällgren J-E (1986) DNA extraction and preliminary characterization of chloroplast DNA from *Pinus sylvestris* and *Pinus contorta*. In: Lindgren D (ed) Provenances and Forest Tree Breeding for High Latitudes. Proceedings of the Frans Kempe Symposium, Umeå, June 10–11, pp 269–280

Szmidt AE, Aldén T, Hällgren J-E (1987) Paternal inheritance of chloroplast DNA in *Larix*. Plant Mol Biol 9:59–64

Tomioka N, Sugiura M (1983) The complete nucleotide sequence of a 16 S ribosomal RNA gene from a blue-green alga, *Anacystis nidulans*. Mol Gen Genet 191:46–50

Wagner DB, Furnier GR, Saghai-Maroof MA, Williams SM, Dancik BP, Allard RW (1987) Chloroplast DNA polymorphisms in lodgepole and jack pine and their hybrids. Proc Natl Acad Sci USA 84:2097–2100

White EE (1986) A method for extraction of chloroplast DNA from conifers. Plant Mol Biol Rep 4:98–101

Yanisch-Perron C, Viera J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119

Communicated by R.G. Herrmann

Received September 30, 1987