

Chloroplast DNA and Molecular Phylogeny

Jeffrey D. Palmer

Summary

The small, relatively constant size and conservative evolution of chloroplast DNA (cpDNA) make it an ideal molecule for tracing the evolutionary history of plant species. At lower taxonomic levels, cpDNA variation is easily and conveniently assayed by comparing restriction patterns and maps, while at higher taxonomic levels, DNA sequencing and inversion analysis are the methods of choice for comparing chloroplast genomes. The study of cpDNA variation has already yielded important new insights into the origin and evolution of many agriculturally important crop plants, and promises to significantly enhance our phylogenetic understanding of the major lines of descent among land plants and algae.

Introduction

The widespread availability of many different restriction endonucleases, together with the development of rapid techniques for cloning, mapping and sequencing DNA, has led to a recent explosion of studies in which DNA variation has been used as the basis for reconstructing the phylogenetic histories of various groups of organisms. The most intense of these efforts has involved the use of mitochondrial DNA sequence variation to describe population structure, introgression, and species relationships among a wide variety of animals, most prominently primates and rodents.^{1,2} Although the pace of DNA systematics research in plants has been slower, there has emerged over the last few years a critical mass of data demonstrating the phylogenetic utility of one of the three plant cellular genomes – the chloroplast genome – for at least one major group of plants, the angiosperms. Before discussing these recent studies, I will first briefly review relevant information on the evolution, inheritance and usefulness in phylogenetic studies of cpDNA as compared with plant mitochondrial and nuclear DNA.

Evolution and Phylogenetic Suitability of cpDNA

Overall, in terms of its size, organization and sequence, cpDNA is the most conservatively evolving genome known.³⁻⁵ Among all characterized land plants, cpDNA has been found to vary less than twofold in size, from 120 to 210 kb.^{4,5} Moreover, almost two-thirds of the observed variation results not from changes in sequence complexity (which varies only from 115 to 150 kb), but from changes in size of a large inverted repeat sequence present in almost all chloroplast genomes. Although chloroplast genomes are very static in size, with large length mutations (i.e. deletions and additions) occurring only very rarely, small length mutations of a few bp to several hundred bp are relatively common during chloroplast genome evolution. The linear order and arrangement of chloroplast sequences is extraordinarily conserved among almost all land plants.^{4,5} With one documented exception (see 'Higher Order Relationships', below), all groups of land plant cpDNAs contain the aforementioned, large inverted repeat sequence (e.g. Fig. 1). Sequence inversions (e.g. Fig. 1) and transpositions occur only very rarely in most chloroplast genomes.

The rate of nucleotide substitutions is extremely low over the chloroplast genome as a whole and in most sequenced genes.^{3,4} In spite of the low overall rate of cpDNA sequence evolution, different chloroplast genes do evolve at different rates,⁴ thus providing a range of molecular yardsticks for measuring evolutionary distances at various taxonomic levels. Accompanying the low rate of nucleotide substitutions in evolution is a relatively low transition-transversion bias.^{3,4} These two factors are reflected in the extremely low incidence of parallel and convergent cpDNA restriction site mutations found in studies at the interspecific level (see next section).

The conservative evolution of the chloroplast genome provides it with distinct advantages in phylogenetic studies compared to the much more dynamic and diverse mitochondria⁴ and nuclear⁶ genomes of plants. Chloroplast genomes are sufficiently small and constrained in size that the entire array of fragments produced by many restriction endonucleases can be easily visualized on a single agarose gel. Such resolution is impossible to obtain for all nuclear genomes and for all but the smallest plant mitochondrial genomes. In addition, the small size of cpDNAs

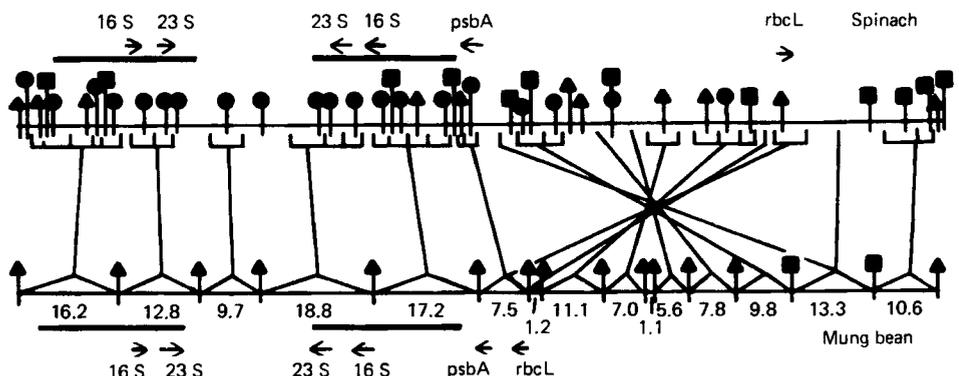


Fig. 1. Mapping of a 50 kb inversion between mung bean and spinach cpDNAs. The circular chloroplast chromosomes are shown linearized to facilitate comparison. Each of 14 cloned mung bean restriction fragments (sizes given in kb) was hybridized to a nitrocellulose filter containing spinach cpDNA restriction fragments. Lines drawn between the two maps indicate regions of homology in the two genomes. Heavy horizontal lines indicate the position of a large inverted repeat in each genome. Reprinted with permission from Cell (ref. 26); copyright held by M.I.T.

REVIEW ARTICLES

(together with their high copy number in leaf cells and relatively straightforward purification) greatly facilitates the isolation and cloning of individual chloroplast genes for sequencing studies. The rapidity with which plant mitochondrial⁴ and nuclear⁶ genomes rearrange their sequences makes it very difficult to describe in any comprehensive and satisfactory manner the evolution of even moderate-sized portions of the genomes, whereas the evolution of an entire chloroplast genome can be described rather easily. Chloroplast genomes are essentially free of any of those evolutionary processes, such as gene duplication and deletion, concerted evolution and pseudogene formation, which are common among nuclear genes and which can dangerously distort the evolutionary history of DNA sequences relative to that of organisms (see discussion in refs. 7 and 8).

Given these disadvantages of mitochondrial DNA relative to the chloroplast genome, plus the lack of any relative advantages (the two genomes have the same pattern of inheritance and similarly slow rates of sequence evolution),⁴ it appears that plant mitochondrial DNA has very little to offer for phylogenetic studies. At best, it is likely to be occasionally useful for purposes of classification, as opposed to phylogenetic reconstruction, as in the case of maize and its relatives.⁹

Although plant nuclear DNA is phylogenetically 'difficult' compared to cpDNA, as a result of its large and variable size, complex modes of rearrangement and concerted evolution, it does possess two attributes that make it quite useful for studies of genetic variation at the interspecific level and below. First, certain nuclear sequences, in particular the ribosomal DNA non-transcribed spacer, appear to evolve more rapidly than any cpDNA sequences. Such nuclear sequences allow a finer level of genetic discrimination, particularly at the population level,¹⁰ than do sequences of the highly conserved chloroplast genome.

Most critically, from the standpoint of phylogeny, nuclear and chloroplast genomes are inherited differently and thus trace potentially different evolutionary histories. The nuclear genome is, of course, always inherited in a biparental, Mendelian fashion, whereas the chloroplast and its genome are often inherited maternally.¹¹ Furthermore, cpDNA inheritance is always clonal, since even where chloroplast inheritance is itself biparental, it is accompanied by somatic segregation in the

complete absence of chloroplast recombination. It is common knowledge that processes of hybridization, introgression and polyploidization are widespread among plants.¹² It is therefore important, whenever taxa capable of hybridization are being studied (generally this means taxa from within a genus), to compare the clonal, generally maternal, phylogeny derived from cpDNA with the biparental phylogeny derived from nuclear DNA and other characters under nuclear control.

Interspecific Relationships

The great majority of studies that have used cpDNA variation to assess phylogenetic relationships have examined cpDNA restriction site variation among closely related plants, usually species within a single genus. Those genera that have been studied in the greatest detail include *Lycopersicon*,¹³ *Pisum*,¹⁴ *Brassica*,^{15, 16} *Triticum-Aegilops*,^{17, 18} *Nicotiana*^{19, 20} and *Solanum*,²¹ while a number of other genera have been

studied in lesser detail (reviewed in refs. 7 and 8).

As a result of the slow rate of cpDNA evolution, restriction patterns of cpDNAs from species belonging to the same genus are often similar enough to allow critical analysis of mutational differences by direct inspection of the gel profiles. For example, the only differences in the Kpn I patterns (Fig. 2) of cpDNAs from eight *Lycopersicon* species and two closely related *Solanum* species are most easily attributed to the presence of one additional Kpn I site in the *Lycopersicon* cpDNAs, resulting in unique fragments of 4.8 kb and 2.8 kb, within the *Solanum*-specific fragment of 7.7 kb. Where more complex enzyme patterns are compared among species at the extremes of a genus's phylogenetic tree it is sometimes necessary to perform specific mapping experiments (usually by filter hybridization) in order to clarify the mutational nature of observed fragment differences.¹⁵

By treating each restriction site as a single two-state character one can use

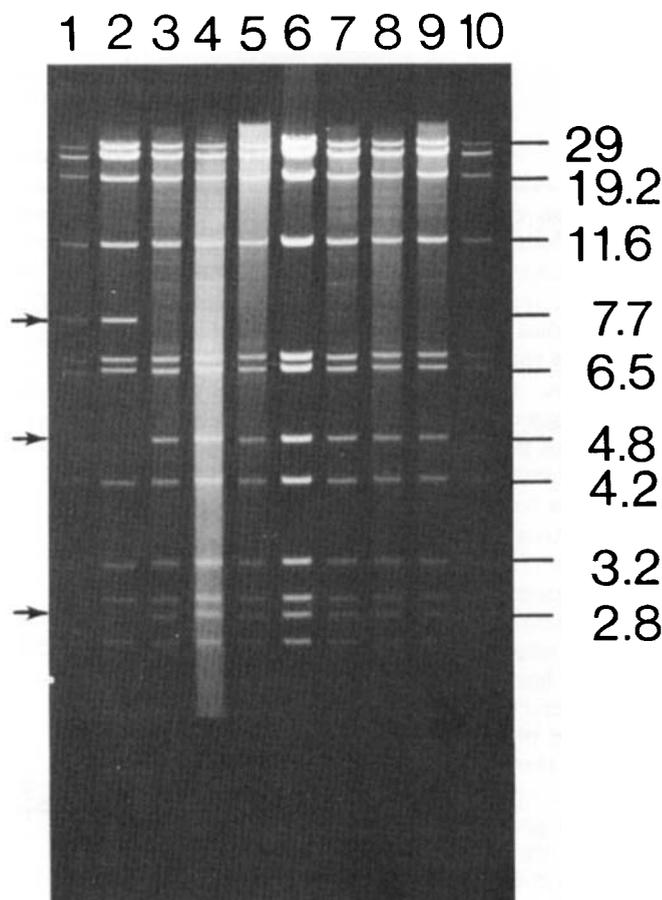


Fig. 2. Restriction site mutation distinguishing eight *Lycopersicon* cpDNAs from two *Solanum* cpDNAs. Purified cpDNAs from ten different species (lane numbers correspond to species accession numbers given in Fig. 3) were digested with *Kpn* I and fragments were separated in a 0.8% agarose gel. Numbers at right are fragment sizes in kb. Arrows indicate fragment size differences ascribable to a specific restriction site mutation (see text). Reprinted with permission from Proceedings of the National Academy of Sciences, USA (ref. 13).

REVIEW ARTICLES

cladistic principles, usually based on the parsimony assumption, to construct a phylogenetic tree which incorporates the evolutionary information from all restriction-site mutations found among a given group of cpDNAs. The tree shown in Fig. 3 is based on analysis of the 10 cpDNAs shown in Fig. 2, plus five other DNAs, with 25 different restriction endonucleases.¹³ Among the 15 DNAs, a total of only 40 independent restriction site mutations (Fig. 3) were detected among 484 restriction sites (2800 base pairs) examined. The very low rate of cpDNA base sequence change is evident from the fact that pairwise sequence divergence values between species ranged from 0% to a maximum of only 0.7%. Moreover, this tree postulates only a single case of homoplasy (convergence or parallelism) in restriction site mutations. This low incidence (2.5%) of homoplasy compares favorably with values of 50% or more that are often obtained in studies using morphological or protein characters, and is one reason why cpDNA holds so much promise for systematic studies.

Relationships within the genus *Lycopersicon* based on the cpDNA-derived phylogeny shown in Fig. 3 are in general agreement with those based on classical criteria, in particular those of morphology²² and crossability.²³ However, the molecular data provide a quantitative estimate of relationships that is hard to obtain otherwise, and in addition provide more detailed resolution at several places. Five separate

restriction site mutations define the three red-orange fruited species in the genus (all other species have green fruits) – *L. esculentum* (the cultivated tomato), *L. pimpinellifolium* and *L. cheesmanii* – as a single monophyletic clade (Fig. 3). This result very strongly suggests that the mutation(s) leading to the ability to synthesize red-orange carotenoid pigments within the tomato fruit occurred only once within this group of species. The cpDNA analysis is in excellent agreement with recent studies^{23,24} which have argued that *S. pennellii* is misplaced in *Solanum* and properly belongs within *Lycopersicon*. The major surprise from the cpDNA study is that the limited amount of intraspecific cpDNA polymorphism found among six accessions of *L. peruvianum* encompasses the two species *L. chilense* and *L. chmielewskii* (Fig. 3). Since *L. chmielewskii* is not thought to be closely related to *L. peruvianum*,²³ this result could indicate either the occurrence of cytoplasmic exchange between the two species, through introgressive hybridization, or else that the degree of their reproductive isolation²³ is not a very good measure of their overall genetic similarity.

A particularly powerful use of cpDNA variation is to elucidate the specific parentage and timing of origin of the hybrid species and polyploid complexes that are so common among plants.¹⁵⁻²¹ For example, two groups of investigators^{15,16} independently analyzed the maternally inherited chloroplast genomes of six *Brassica* species in order to elucidate the maternal and paternal parentage of the three amphidiploid species whose putative nuclear origins are shown in Fig. 4. Both groups were able to assign unambiguously the maternal progenitors for the amphidiploids *B. carinata* and *B. juncea* as *B. nigra* and *B. campestris*, respectively (and by subtraction, that *B. oleracea* and *B. nigra* must have served as the paternal parents in these crosses). In addition, Palmer *et al.*¹⁵ were able to identify a specific line among several different lines of each of *nigra* and of *campestris* that is most likely to have served as the maternal progenitor for each of the two diploids.

Quantitative analysis of the number of cpDNA restriction site mutations distinguishing a given amphidiploid from each diploid parent, and also the two parents from each other, allowed insight into the relative timing of the interspecific hybridization event. Thus, the fact that the *carinata* chloroplast genome is identical at all 3000 bp of the restriction sites compared to that of a

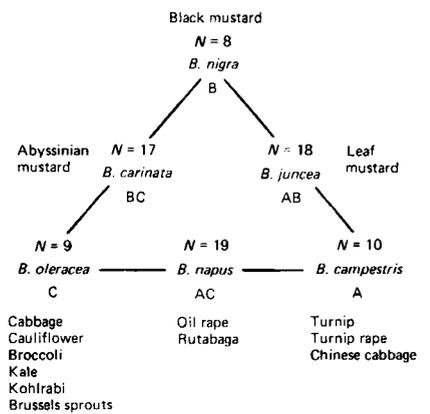


Fig. 4. Hypothesized origins of three amphidiploid species of *Brassica* from three diploid species. Haploid chromosome numbers and genome designations are indicated. Based on cytogenetic studies of U.²⁵

specific *nigra* line allowed the conclusion that this hybridization must be a very recent event.¹⁵ The fact that *nigra* and *oleracea* differ at a large number of restriction sites (2.4% sequence divergence) indicates that the two species must have diverged quite some time ago relative to their recent hybridization to form *carinata*.¹⁵

Most surprisingly, both studies^{15,16} found that the chloroplast genome of the amphidiploid *B. napus* is quite unlike that of either diploid species (Fig. 4) that gave rise to the *napus* nucleus. In fact, formal cladistic analysis indicates that the *napus* cytoplasm may actually have diverged from a cytoplasm which is a common ancestor of the *oleracea* and *campestris* cytoplasm, i.e. that the *napus* cytoplasm divergence predates that of the two diploids from one another.¹⁵ To explain this discordancy between maternal, chloroplast and biparental, nuclear phylogenies, it was postulated that *B. napus* received its cytoplasm by introgression of some foreign, unidentified *Brassica* cytoplasm.¹⁵ Thus the cpDNA results suggest that *B. napus* has a particularly complicated history, involving both interspecific hybridization, as well documented between *B. oleracea* and *B. campestris*, and possibly also introgressive hybridization.

Higher Order Relationships

Compared to the wealth of phylogenetic data gathered from cpDNA systematic studies at the interspecific level, relatively little effort has been made to apply molecular techniques to the study of higher-order relationships among plant families, orders and classes. Four studies have extended the approach of compa-

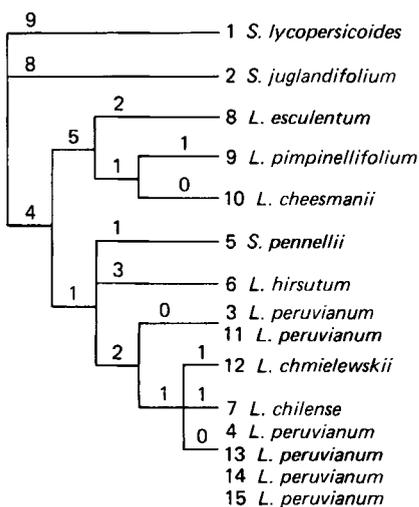


Fig. 3. CpDNA phylogeny for *Lycopersicon* and two related species in *Solanum*. Numbers at termination of branches indicate accessions (see Table 1 of ref. 13). Numbers on the branches indicate the number of cpDNA restriction site mutations specific to each branch. Reprinted with permission from Proceedings of the National Academy of Sciences, USA (ref. 13).

REVIEW ARTICLES

ring restriction fragment patterns and maps in order to study evolutionary relationships among different genera within a family (reviewed in refs. 7 and 8). Unfortunately, the phylogenetic insights gained in these studies have been necessarily limited by the small numbers of taxa examined.

As mentioned in an earlier section, major structural rearrangements, i.e. large length mutations, inversions and transpositions, occur very rarely during the course of chloroplast genome evolution. However, when found, these mutations often serve as very prominent and powerful phylogenetic markers, demarcating major dichotomies among plant groups. For example, the large inverted repeat sequence shown in Fig. 1 for the chloroplast genomes of spinach and mung bean is an evolutionarily very ancient structure that has been lost only once among all land plants examined. One entire segment of this duplicated sequence has been deleted from the chloroplast genomes of species belonging to at least eight genera and four tribes of legumes, while the repeat structure has been retained within species (including mung bean) from at least four other genera and two other tribes of legumes.^{7, 8, 26} Thus this deletion mutation makes a deep and profound split within the legume family. The 50 kb sequence inversion shown in Fig. 1 is shared by the chloroplast genomes of all twelve genera and six tribes of the legume subfamily Papilionoideae that have been examined.^{7, 8, 26} It will be interesting to see whether this inversion is also shared by species belonging to the other two legume subfamilies (Mimosoideae and Caesalpinioideae), and perhaps also by related families in the subclass Rosidae. Several other cpDNA sequence inversions have also been found that make major phylogenetic groupings among and within families of flowering plants (reviewed in refs. 7 and 8).

A particularly profound phylogenetic insight gained by comparing the sequence of genes in chloroplasts and in eubacteria, cyanobacteria in particular, is that chloroplasts almost certainly originated by the endosymbiotic association of a photosynthetic prokaryote and a primitive protoeukaryote.^{4, 27, 28} Furthermore, sequence comparisons suggest that different plastid lineages were established by multiple independent endosymbioses.

Conclusions

By virtue of its easy purification, small size, highly constrained organization, and slow rate of sequence evolution, cpDNA is particularly well suited for phylogenetic studies at a range of taxonomic levels. Thus far, this potential has been demonstrated primarily in interspecific studies, many of which have examined the origin and evolutionary relationships of various important crop plants. A feature which has provided unique insights into the parentage and timing of origin of interspecific hybrid species and polyploid complexes is the predominantly maternal mode of inheritance of cpDNA. Although little effort has been made to apply cpDNA analysis to resolving higher-order taxonomic relationships, it is predicted that significant advances in our understanding of the evolutionary relationships of major groups of plants will come primarily from comparative sequence studies of chloroplast genes.

I thank Dr R. K. Jansen for critical reading of the manuscript. This work was supported in part by a grant from the National Science Foundation (BSR-84-15934).

REFERENCES

- 1 AVISE, J. C. & LANSMAN, R. A. (1983). Polymorphism of mitochondrial DNA in populations of higher animals. In *Evolution of Genes and Proteins* (eds. M. Nei & R. K. Koehn), pp. 147-164. Sinauer, Sunderland, Mass.
- 2 BARTON, N. & JONES, J. S. (1983). Mitochondrial DNA: new clues about evolution. *Nature* **306**, 317-318.
- 3 CLEGG, M. T. & CURTIS, S. E. (1984). Molecular evolution of chloroplast DNA sequences. *Mol. Biol. Evol.* **1**, 291-301.
- 4 PALMER, J. D. (1985). Evolution of chloroplast and mitochondrial DNA in plants and algae. In *Monographs in Evolutionary Biology: Molecular Evolutionary Genetics* (ed. R. J. MacIntyre). Plenum, New York (In the press.)
- 5 PALMER, J. D. (1985). Comparative organization of chloroplast genomes. *Annu. Rev. Genet.* **19** (In the press.)
- 6 FLAVELL, R. B. (1980). The molecular characterization and organization of plant chromosomal DNA sequences. *Annu. Rev. Plant Physiol.* **31**, 569-596.
- 7 PALMER, J. D. (1985). Phylogenetic analysis of chloroplast DNA variation. *Ann. Missouri Bot. Garden* (In the press.)
- 8 PALMER, J. D. (1985). Chloroplast DNA and phylogenetic relationships. In *DNA Systematics: Evolution* (ed. S. K. Dutta). CRC Press, Boca Raton, Florida (In the Press.)
- 9 WEISSINGER, A. K., TIMOTHY, D. H., LEVINGS, C. S. III & GOODMAN, M. M. (1983). Patterns of mitochondrial DNA variation in indigenous maize races of Latin America. *Genetics* **104**, 365-379.
- 10 SAGHAI-MAROOF, M. A., SOLIMAN, K. M., JORGENSEN, R. A. & ALLARD, R. W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. USA* **81**, 8014-8018.
- 11 SEARS, B. B. (1980). The elimination of plastids during spermatogenesis and fertilization in the plant kingdom. *Plasmid* **4**, 233-255.
- 12 GRANT, V. (1981). *Plant Speciation*. Columbia University Press, New York.
- 13 PALMER, J. D. & ZAMIR, D. (1982). Chloroplast DNA evolution and phylogenetic relationships in *Lycopersicon*. *Proc. Natl. Acad. Sci. USA* **79**, 5006-5010.
- 14 PALMER, J. D., JORGENSEN, R. A. & THOMPSON, W. F. (1985). Chloroplast DNA variation and evolution in *Pisum*: patterns of change and phylogenetic analysis. *Genetics* **109**, 195-213.
- 15 PALMER, J. D., SHIELDS, C. R., COHEN, D. B. & ORTON, T. J. (1983). Chloroplast DNA evolution and the origin of amphidiploid *Brassica* species. *Theor. Appl. Genet.* **65**, 181-189.
- 16 ERICKSON, L. R., STRAUS, N. A. & BEVERSDORF, W. D. (1983). Restriction patterns reveal origin of chloroplast genomes in *Brassica*. *Theor. Appl. Genet.* **65**, 201-206.
- 17 BOWMAN, C. M., BONNARD, G. & DYER, T. A. (1983). Chloroplast DNA variation between species of *Triticum* and *Aegilops*. Location of the variation on the chloroplast genome and its relevance to the inheritance and classification of the cytoplasm. *Theor. Appl. Genet.* **65**, 247-262.
- 18 TERACHI, T., OGIHARA, Y. & TSUNEWAKI, K. (1984). The molecular basis of genetic diversity among cytoplasm of *Triticum* and *Aegilops*. III. Chloroplast genomes of the M and modified M genome-carrying species. *Genetics* **108**, 681-695.
- 19 KUNG, S. D., ZHU, Y. S. & SHEN, G. F. (1982). *Nicotiana* chloroplast genome. III. Chloroplast DNA evolution. *Theor. Appl. Genet.* **61**, 73-79.
- 20 SALTS, Y., HERRMANN, R. G., PELEG, N., LAVI, U., IZHAR, S., FRANKEL, R. & BECKMAN, J. S. (1984). Physical mapping of plastid DNA variation among eleven *Nicotiana* species. *Theor. Appl. Genet.* **69**, 1-14.
- 21 HOSAKA, K., OGIHARA, Y., MATSUBAYASKI, M. & TSUNEWAKI, K. (1984). Phylogenetic relationships between the tuberous *Solanum* species as revealed by restriction endonuclease analysis of chloroplast DNA. *Jpn. J. Genet.* **59**, 349-369.
- 22 MULLER, C. H. (1940). A revision of the genus *Lycopersicon*. *U.S. Dept. Agric. Misc. Publ.* **382**.
- 23 RICK, C. M. (1979). Biosystematic studies in *Lycopersicon* and closely related species of *Solanum*. In *The Biology and Taxonomy of the Solanaceae*, *Linnean Soci-*

ety Symposium Series No. 7 (eds. J. W. Hawkes, R. N. Leash & A. D. Skelding), pp. 667–678. Academic Press, New York.

24 D'ARCY, W. G. (1982). Combinations in *Lycopersicon* (Solanaceae). *Phytologia* 51, 240.

25 U, N. (1935). Genomic analysis in *Brassica* with special reference to the experimental formation of *B. napus* and

peculiar mode of fertilization. *Jpn. J. Bot.* 7, 389–452.

26 PALMER, J. D. & THOMPSON, W. F. (1982). Chloroplast DNA rearrangements are more frequent when a large inverted repeat is lost. *Cell* 29, 537–550.

27 GRAY, M. W. (1983). The bacterial ancestry of plastids and mitochondria. *Bioscience* 33, 693–699.

28 GRAY, M. W. & DOOLITTLE, W. F. (1982). Has the endosymbiont hypothesis been proven? *Microbiol. Rev.* 46, 1–42.

JEFFREY D. PALMER is in the Division of Biological Sciences, University of Michigan, Ann Arbor, Michigan 48109, USA.

The Astral Relaxation Theory of Cytokinesis Revisited

J. G. White

Summary

Cytokinesis in animal cells is accomplished by the active constriction of the equatorial regions of a cell by an actomyosin-containing contractile ring. The mitotic apparatus specifies the position and orientation of the furrow such that the mitotic spindle is always bisected. Global cortical contractions occur in the cortex of a cell prior to cytokinesis that are independent of the presence of the mitotic apparatus. It was proposed some years ago that the asters of the mitotic apparatus could act to relax the preformed cortical tension in their vicinity. This would produce a differential in tension between the equatorial regions and the adjacent regions of the cortex so that the equatorial regions would contract, forming a cleavage furrow. It can be shown that, as it stands, this theory cannot explain cleavage. However, if cortical contractile elements are assumed to be laterally mobile in the plane of the cortex, then the astral relaxation theory can account for many of the aspects of cleavage, including the formation of the contractile ring. Similar schemes may account for the behaviour of the lamellipodia of motile cells.

Introduction

Cytokinesis and mitosis are the fundamental mechanisms that bring about the division of cells. The two mechanisms work in concert, the cleavage furrow always bisecting the mitotic apparatus (MA) after the chromosomes have separated and have migrated to the poles, thus ensuring the equipartition of genetic material to the daughter cells. In spite of the fundamental nature of

cytokinesis, it has received relatively scant attention from contemporary cell biologists. The subject has, however, been studied for many years by a few enthusiasts, who have accumulated a considerable body of observational and experimental data. A knowledge of the mechanisms involved in cytokinesis will have considerable intrinsic interest because of the ubiquity of this process, and

will also probably shed light on other aspects of cellular motility, such as cell locomotion and morphogenetic movements during development. I will attempt to review briefly the current state of knowledge on the subject and describe a model that was proposed some time ago in the light of some recent observations.

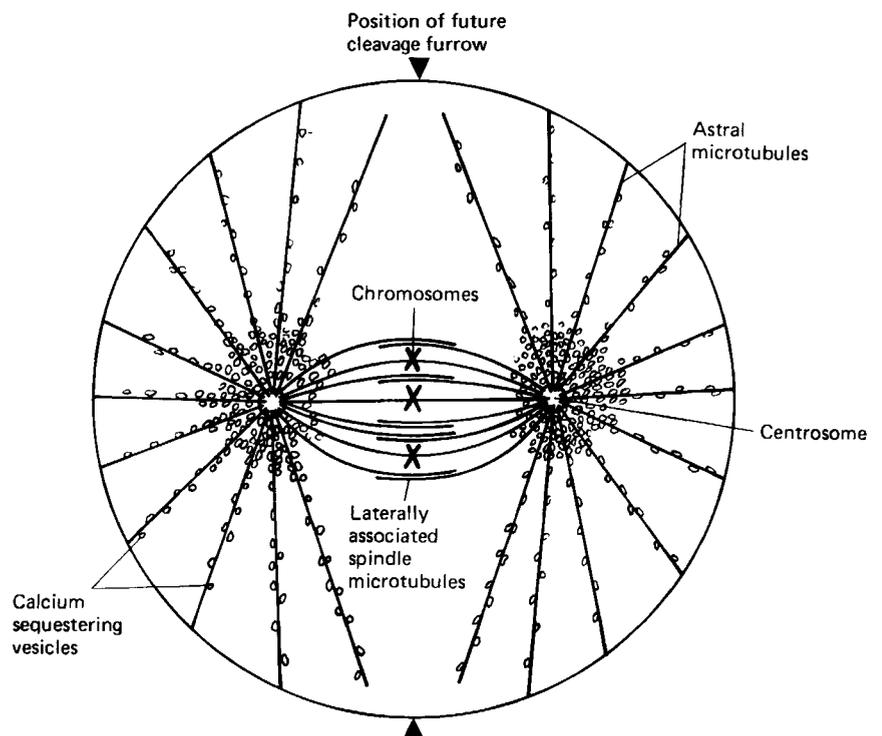


Fig. 1. The mitotic apparatus of a typical animal cell at metaphase. Microtubules radiate out of two centrosomes. Some are attached to the kinetochores of chromosomes and others become laterally associated with microtubules from the opposite aster, making up the body of the mitotic spindle. Astral microtubules away from the spindle region radiate out to the cortex of the cell. Large accumulations of membrane-bound vesicles are associated with the astral microtubules, some of which probably have a calcium-sequestering activity. Furrowing occurs at the equator (arrowed) after the chromosomes have separated and migrated to the poles.