The Sexual Nature of the Eukaryote Genome

G. Bell

This paper supports a previous conjecture that the sexual cycle of eukaryotes arose from the infection of cells by genome parasites. The finding are as follows. (1) In prokaryotes, conjugative plasmids ensure their own spread by directing partial cell fusion. (2) Conjugative plasmids permit gene transfer between prokaryotes and eukaryotes, and can be integrated into the eukaryote genome. (3) Genes can be transferred between unrelated eukaryotes, and between different genomes within eukaryotic cells. (4) Elements such as transposons and retroviruses evolve as parasites of the eukaryote sexual system. (5) The mating-type genes of bipolar fungi are idiomorphic: alternative genes directing sexual specificity are dissimilar and nonhomologous. It is argued that they arose as parasitic elements directing cell fusion, which have become integrated into the genome. (6) Mating-type idiomorphs in different taxa may be dissimilar, reflecting the independent acquisition of different infectious elements by different eukaryote lineages. (7) Sexual competence is rapidly lost through mutation accumulation or antagonistic pleiotropy during vegetative proliferation, so that many lineages are sexually sterile. (8) In multipolar systems, novel mating-type alleles arising by mutation spread in a parasite-like manner by virtue of their access to lineages which have become sexually sterile. (9) It is suggested that centromeres arise as devices to enable low copy number plasmids to persist. (10) Crossing over is favored if it enables the genes that direct it to unlink themselves from inferior genomes, or from the consequences of their own transposition. (11) Elements that are originally infectious—and which by directing cell fusion, regular segregation, or crossing over promote their own replication-become domesticated by vertical transmission when sex becomes associated with reproduction, thereby evolving into the characteristic features of the eukaryote genome.

The life cycle of eukaryotes comprises two distinct cycles: a vegetative cycle of growth and reproduction, and a sexual cycle of fusion and restitution (see Bell and Koufopanou 1991). In large multicellular forms, such as plants and animals, these cycles are usually conflated because of the obligate association of sex with reproduction. In unicellular, coenobial, and mycelial eukaryotes, the two cycles are separate and their relationship can be studied more easily. In this article, I argue that the characteristic features of the eukarvote genome arise from the entrainment of parasitic self-replicating genetic elements into the life cycle. The argument is based on three well-known theories:

1. The original nature of the eukaryotic cell as a community of endosymbiotic prokaryotes, whose independence can still be inferred from their retention of autonomously replicating genomes (see Margulis 1981). 2. The nature of genetic elements such as transposons and some introns as selfreplicators, not contributing to the harmonious function of the individual cell or organism, which arise as horizontally transmitted infections, and which could themselves have been the genes originally responsible for the establishment of a sexual cycle (Hickey 1982; Hickey and Rose 1988; Rose 1983).

3. The tendency for parasitic elements, or organisms, originally transmitted horizontally, to be domesticated into a mutualistic relationship with the rest of the genome by vertical transfer.

I shall use recent analyses of the molecular structure of genomic elements to generalize these principles and to interpret the organization of the eukaryote genome as a community of originally parasitic genetic elements which have become in part mutualistic through their association in vertical transmission.

Journal of Heredity 1993;84:351-359; 0022-1503/93/\$5.00

From the Biology Department, McGill University, 1205 Avenue Dr., Penfield, Montreal, Quebec, Canada H3A 1B1. This paper was delivered at a symposium titled "The Evolution of Sex" sponsored by the American Genetic Association at Virginia Polytechnic Institute and State University in Blacksburg, Virginia, on July 10 and 11, 1992.

The Infectious Nature of Bacterial Sexuality

In bacteria, the parasitic nature of conjugation has long been apparent and is well understood mechanistically at a molecular level. Cells bearing an F plasmid transmit a copy of this plasmid to an uninfected partner, through the ability of the plasmid to encode partial cell fusion and gene transmission. Conjugation between two infected cells is inhibited by cell-surface proteins encoded by the F plasmid. If the plasmid is integrated into the host genome it remains infectious, but conjugation may now result in the transfer of a part of the host genome to the recipient cell, as well as, or instead of, the plasmid itself. The horizontal transmission of nonplasmid genes may have important consequences for bacterial populations, but the evolution of conjugative plasmids can be understood very straightforwardly in terms of selection for elements that are able to spread to a range of lineages, rather than being bound to a single lineage which will sooner or later become extinct, without invoking any more onerous populationlevel process. The evolution of sexual specificity, as the preferential infection of previously uninfected lineages, is an important consequence of selection for efficient transfer.

Similar processes, in which interaction between an infected and an uninfected cell results in two infected cells, are known from eukaryotic microbes. For example, in Chlamydomonas reinhardtii the chloroplast genome is usually transmitted only by the mt+ gamete; following sexual fusion, the chloroplast genome of its mtpartner is destroyed by endonucleases, and all four products of meiosis bear the chloroplast genome of their mt+ parent. The mitochondrial genome is transmitted uniparentally in the opposite direction. However, the mating-type genes which direct sexual fusion are present in the nuclear and not in the chloroplast genome. A closer approach to the bacterial system is made by plastid genes which encode their own proliferation after cell fusion. An intron in the rRNA-coding chloroplast gene of C. eugametos encodes a site-specific endonuclease which cuts near the homologous intronless site in the closely related form C. moewusii, where a copy of the intron can then be inserted, so that all products of the cross carry the intron (Marshall and Lemieux 1991). The apocytochrome gene carried on the mitochondrial genome of C. smithii is colinear with the homologous gene in *C. reinhardtii*, except for a 1-kbp insert (the " α insert") (Boynton et al. 1987). This insert also behaves like a mobile intron which is transmitted to all sexual products (Remacle et al. 1990), perhaps again through encoding a site-specific endonuclease (Colleaux et al. 1990). The ω intron of yeast mitochondria behaves in a similar way (Jacquier and Dujon 1985).

Cytological and genetic evidence shows that mitochondria in yeast and *Chlamydomonas* can fuse and produce recombinant plastids. Elements such as α in *Chlamydomonas* and ω in yeast are parasites that exploit this cycle of fusion and separation. Mitochondria in the plasmodial slime-mould *Physarum polycephalum* carry a plasmid which is not only parasitic, being transmitted to all descendants, but which also itself promotes fusion with other mitochondria (Kawano et al. 1991). This represents a genetic element that spreads by virtue of its ability to initiate a sexual process.

When eukaryotic cells first appeared, their component prokaryotes must have carried with them their complement of genetic parasites, including mobile elements and conjugative plasmids. It is now becoming clear that these elements can still be identified in the plastids of modern forms, where they exploit and induce processes analogous to sexual fusion, just as they do in freeliving bacteria.

Occasional Genetic Transfer Between Distant Lineages

Genes are exchanged regularly among prokaryote taxa under natural conditions (Stotzky and Babich 1986; Triess-Cuot et al. 1988). They may also, though much less frequently, be transferred from prokaryotes to eukaryotes. The mob system of the broad-host-range plasmid RSF1010 enables transfer not only among gram-negative bacteria but also from Agrobacterium to plant cells (see Zambryski 1988). Extensive prokaryote-to-prokaryote transfer may imply that prokaryote-to-eukaryote transfer is achieved relatively easily by the same mechanisms. Agrobacterium is a specialized parasite, but the possibility that similar processes may occur rather generally has been raised by the discovery of conjugal transfer of genes from Escherichia coli to yeast (Heinemann and Sprague 1989). This very wide transfer is mediated by conjugative plasmids which would probably not themselves be maintained stably in a yeast lineage. However, such plasmids occasionally become integrated into a yeast chromosome, and they are then stably replicated and maintained (Nishikawa et al. 1992). The acquisition and subsequent integration of prokaryotic conjugative plasmids by eukaryotic microbes seems to be a process readily observed and manipulated in the laboratory, although its occurrence in natural populations has not yet been investigated.

The transfer of retroviral sequences among different species of animals seems to be rather commonplace (Benveniste 1985). There is some evidence that transposable elements can be transmitted between species of Drosophila by processes other than normal mating. Possible examples include the P-element, which encodes a transposase, and seems to have been acquired by D. melanogaster from D. willistoni (see Daniels et al. 1990), and the jockey element, which includes an internal promoter and encodes a reverse transcriptase, and seems to have been acquired by D. melanogaster from D. funebris (Mizrokhi and Mazo 1990). These three Drosophila species are classified in different sections within the genus, and vertical transmission from a common ancestor seems unlikely because the elements are missing from more closely related species within the same group. It has been reported recently that the mite Proctolaelaps can transmit P elements between hosts while feeding, and may therefore be responsible for transferring them between species.

Genetic transmission from eukarvotes to prokaryotes is conceivable, but has not yet been firmly established. Martin and Fridovich (1981) proposed that the symbiotic bacterium Photobacter leiognathi had acquired a superoxide dismutase gene from its host, the ponyfish Leiognathus splendens. This inference has been criticized and defended by a number of authors (Bannister and Parker 1985; Cornish-Borden 1985; Leunissen and de Jong 1986), and remains controversial. Doolittle et al. (1990) present a detailed phylogenetic argument for the transfer of a glyceraldehyde-3-phosphate dehydrogenase gene from an unknown eukaryote to E. coli. The only known bacterial hemoglobin, from Vitreoscilla, is more similar to plant leghemoglobins than to animal hemoglobins, and may have been acquired from a vascular plant host (Wakabayashi et al. 1986). Carlson and Chelm (1986) suggested that the root-nodule bacterium Rhizobium had acquired a glutamine synthetase gene from its plant host, but this has been disputed (Shatters and Kahn 1989).

Chloroplasts and mitochondria originated as endosymbiotic bacteria, but their genomes lack most of the genes that would be required for freeliving existence. It is therefore a consequence of the theory of serial endosymbiosis that the missing genes have been either lost entirely or transferred to the nuclear genome. Plastid genes can be transferred artificially to the nuclear genome (Nagley and Devenish 1989). The most convincing example of a natural transfer is the gene encoding chloroplast protein elongation factor, which is found in the chloroplasts of green algae but in the nucleus of vascular plants (Baldauf and Palmer 1990). The transfer of genes from chloroplasts to mitochondria may be quite frequent and has given rise to the term "promiscuous DNA" (Farrelly and Butow 1983; Stern and Palmer 1984). The analogous process in prokaryotes is the integration of plasmid genes into the major genome. In Streptococcus, conjugation can be mediated by chromosomally located transposons rather than by autonomous plasmids (Clewell and Gawron-Burke 1986).

It seems likely, therefore, that eukaryote lineages are infected from time to time by exotic genetic elements, which once stabilized can move from genome to genome within the cell.

Transposons and Retroviruses Are Common Parasites of the Eukaryote Sexual Cycle

A large fraction of many eukaryote genomes does not code for protein, or at least not for proteins that contribute to function at the level of the cell. These sequences persist because they exploit the replicative machinery of the cell or are able to encode their own self-replication. Some have extended their parasitic capabilities by being able to proliferate copies within the genome by a variety of mechanisms. In strictly asexual lineages, vertically transmitted parasites such as transposons, retroviruses, and B-chromosomes cannot accumulate to any great extent, because any asexual lineage is certain eventually to become extinct, and when it does so all its genetic parasites become extinct too. If these parasites are very effective in spreading through the genome they are likely to damage the cell, or the individual, and will thus hasten the extinction of the clone. Vertically transmitted parasites of asexual organisms are therefore likely to be infrequent and benign. These restrictions do not apply to

sexual lineages, in which genetic parasites can infect new lines of descent through sexual fusion. High rates of self-replication can then be favored even when high copy number is deleterious to the host individual, because the success of the parasite is not indissolubly linked to the survival of a clonal lineage or host individual. Sex favors the evolution of virulent, rapidly spreading genetic parasites.

The formal population genetics of this process have been worked out (see Charlesworth and Langley 1986), but as yet there is little experimental or comparative support for these arguments. It is not known whether transposition rates or copy number are typically higher in organisms with more frequent sexual episodes, although B chromosomes are more frequent among outcrossing than among inbreeding plants (see Bell and Burt 1990). Copy number of the Gulliver transposon of C. reinhardtii did not decrease during 400 generations of asexual growth in batch culture, although transposition rate was not measured (Zeyl C and Bell G, unpublished). Among obligately sexual organisms, transposition rates should be greater in species with more chromosomes, but there is not yet sufficient evidence to test this prediction. The evolutionary biology of molecular parasites is still an almost open field for empirical investigations (see also Maynard Smith and Nee 1989).

Mating-Type Genes as Ancient Infections

The most remarkable conclusion to emerge so far from investigations of the molecular structure of mating-type genes in fungi is that alternative genes are often not allelic. Instead of differing with respect to a few bases or rearrangements, they are generally completely dissimilar. The term "idiomorph" has been coined to denote alternative genes occupying homologous positions on a chromosome but having sequences so dissimilar that the genes themselves are unlikely to be homologous (Metzenberg and Glass 1990). Neurospora crassa is a filamentous bipolar ascomycete with mating types designated A and a. These genes control both sexual and vegetative compatibility: sexual fusion occurs only between mycelia with different genes, and successful vegetative fusion occurs only between mycelia with similar genes. The nuclei in each mycelium carry a single copy of one of the mating-type genes. Transformation experiments have shown that differences at the mating-type locus are alone necessary and sufficient to determine compatibility. The two mating-type genes are idiomorphs. The A idiomorph comprises a stretch of 5,301 bp, which includes a single open reading frame of 928 bp containing a single intron and encoding a polypeptide of 288 amino acids which is responsible for sexual and vegetative compatibility (Glass et al. 1990). The a idiomorph comprises 3,235 bp, of which a 1,260-bp sequence containing two introns encodes a polypeptide of 382 amino acids and is again responsible for sexual and vegetative compatibility (Staben and Yanofsky 1990). The open reading frames in the two idiomorphs are located at opposite ends of the mating-type region and are transcribed in opposite directions. It is difficult to believe that these two genes have diverged by mutation from a common ancestral sequence (Metzenberg and Glass 1990); they must then have been acquired independently.

Podospora anserina is a second filamentous ascomycete with bipolar idiomorphic mating-type genes. One mating type has been subcloned to a 22-kbp fragment, while the other is represented by two fragments of 18.0 and 4.3 kbp (Picard et al. 1991).

The unicellular ascomycete Saccharomyces cerevisiae is also bipolar, with mating types designated a and α , but each cell bears a copy or copies of both mating-type genes. One copy of each gene is silent; a third locus, situated between the two storage loci, bears a copy of one of the two genes, which is expressed and thereby determines the sexual phenotype of the cell, but which is frequently replaced after reproduction by the alternative gene. Gender is determined by dissimilar sequences of 642 bp in MATa and 747 bp in MATa, these being flanked by sequences that are very similar in the two mating types. A similar bipolar system for switching gender is found in Scizosaccharomyces pombe, where mating type is determined by two idiomorphic genes each about 1 kbp in length. Gender determination in yeasts has been reviewed by Herskowitz (1988), Kelly et al. (1988), and Egel et al. (1990). The transposition of sequences from silent storage sites to the expression site, displacing the resident sequence, is strikingly reminiscent of the behavior of genetic parasites such as transposons. It is also interesting that a mating-type gene, together with some of the intervening DNA, can loop out spontaneously from the chromosome, forming a small circular plasmid (Beach and Klar 1984). Such plasmids can later reintegrate at the correct site, as shown

by their ability to cure otherwise lethal deletions of the corresponding region of the chromosome (Heim and Gutz 1991).

The basidiomycete Schizophyllum commune has a more complex system of gender determination in which the A locus determines the mating reaction while the B locus affects the subsequent development of viable products. Both loci are in turn complex: the A locus is in fact two linked loci, $A\alpha$ and $A\beta$, which have nine and 32 known states, respectively. Two mycelia will mate only if they bear different genes at $A\alpha$ or $A\beta$ or both; nuclei bear single copies of the genes, and no switching of gender occurs. Giasson et al. (1989) have studied two of the $A\alpha$ genes and have found them to be idiomorphic. $A\alpha I$ was subcloned within a 2.8-kbp fragment and $A\alpha 4$ within a 1.2-kbp fragment. The two fragments have dissimilar restriction maps, and probes containing one will not bind to DNA containing the second sequence, or to genes at the $A\beta$ or B loci.

Coprinus cinereus has a similar but perhaps even more complex array of loci determining sexual specificity. The A42 mating-type factor comprises two very closely linked genes, $A\alpha 3$ and $A\beta 3$. Fragments containing these two genes did not hybridize, nor did A42 DNA hybridize with the A5 and A6 mating-type factors; consequently, Mutasa et al. (1990) identified the mating-type genes as idiomorphs. May et al. (1991) found that the A43 matingtype factor comprised three dissimilar subunits. When these subunits were separately cloned, they found extensive but very variable hybridization with other strains carrying different A factors. The situation is clearly very complicated; it seems possible that variants of each constituent A unit belong to several groups, members of the same group being to some extent similar, and possibly allelic, while members of different groups are completely dissimilar idiomorphs.

In Ustilago maydis the A locus is bipolar and the B locus is multipolar, but both are relatively simple in structure. The two A genes, which are alone responsible for sexual specificity, have been cloned and are idiomorphs of different sizes (4-5 versus 7.5-8.5 kbp) and with dissimilar restriction maps (Froeliger and Leong 1991). The B locus, which directs development after fusion, has a long series of 20-30 different states. The six B genes that have been studied are very similar and clearly allelic, each with an open reading frame of the same size and limited variability in sequence (Kronstad and Leong 1990). The mating-type genes of *Gelasinospora* and *Cochliobolus* are also said to be idiomorphic (Metzenberg 1990, citing unpublished results).

These patterns are difficult to understand in terms of conventional theories in which alternative mating-type alleles arise by mutations, rearrangements, or duplications and serve merely as markers to regulate outcrossing. It is conceivable that they have arisen from systems in which fusion was regulated by alternative alleles at each of two tightly linked loci, by deletion of one allele at each locus. However, it is then very surprising that all the bipolar fungi so far investigated have turned out to be idiomorphic, since retaining an allelic system must be more likely than undergoing a complementary double deletion. In Schizophyllum fusion is regulated by two linked loci, but the two $A\alpha$ genes so far sequenced are idiomorphic; to explain this requires a system consisting originally of three loci. Nor is it clear why genes which regulate fusion should so frequently have undergone a double deletion, while those which regulate subsequent development have, so far as we know, never done so. The observations may be easier to understand if eukaryotic sex evolved from a directly infectious system such as the conjugative plasmids of bacteria. Imagine that such a plasmid, with the property of directing cell fusion, succeeded in invading a eukaryote lineage and became integrated into the chromosomal genome, as an element akin to a retrovirus or a transposon. It might be a single gene, encoding a single polypeptide, or a small set of tightly linked genes. Because of its ability to infect related asexual lineagesconverting them to sexuality in the process-it will rapidly spread through a community, even if it reduces vegetative fitness. As in the case of bacterial plasmids, selection will favor autoimmunity, so that infected cells will not fuse with one another. If the element were to spread to fixation, within the community of microbes susceptible to it, fusion will cease, since virtually all cells are infected, and we shall observe an apparently asexual taxon. Mutation might then degrade the element to the point where it would be unrecognizable except as a stretch of unique but noncoding DNA. However, it might still be possible to recognize such elements, and even elicit their expression, in perennially asexual taxa such as euglenids or rhizopods. A second element with similar properties could meanwhile sweep through the population. If two such elements were segregating simultaneously, their interests would conflict, because potential sexual partners would often both be infected by one element while only one was infected by the other. This conflict would be resolved if the elements came to occupy homologous positions on a chromosome, which is not an unreasonable speculation, given that each originated as something like a transposon, able to move within the genome. The outcome would then be a stable sexual system with the embedded parasites acting as alternative mating types.

Though highly speculative, this scheme is compatible with many features of the molecular structure of mating-type genes. First, the mating-type genes of fungi seem generally to be idiomorphic, not allelic. Mutation does not create genes with novel mating specificities, but only defective or inactive genes. The most reasonable explanation of idiomorphism is that alternative mating-type genes have descended from different ancestral sequences. Second, in systems in which sexual and vegetative compatibilities are separated, the genes that control fusion are idiomorphic whereas those that control subsequent development are allelic. Idiomorphism is therefore specifically a feature of systems involved in sexual fusion and is presumably related to the evolutionary dynamics of elements that direct fusion. In strictly bipolar forms in which sexual and vegetative compatibility are controlled at the same locus, sexual fusion requires that partners possess different idiomorphs while somatic fusion requires that they both possess the same idiomorph. This pattern is readily explained on the basis of an interpretation of idiomorphs as parasites: they should attempt to infect lineages from which they were previously absent but to pool cytoplasmic resources with individuals who are already infected. Third, in yeasts the idiomorphs that direct fusion function by transposition, which can be interpreted as a modification of their original status as transposons.

Elements able to direct cell fusion will do so by encoding, regulating, or being closely linked to genes involved in cell signaling: in *Ustilago*, for example, the mating-type region contains open reading frames that may encode a pheromone and its receptor (Bölker et al. 1992). It would be equally unsurprising to discover that the sequences responsible for encoding mating-type switching in yeasts bore some resemblance to transposons. The crucial observation would be that sequences within the now untranscribed and apparently functionless section of the mating-type region in organisms such as *Neurospora* and *Chlamydomonas* bore a detectable resemblance to transposons, since, in the absence of present function, such a resemblance could be ascribed only to ancestry.

If mating-type genes originated as infections, then different eukaryote lineages are likely to have been infected by different parasitic sequences. Genes that control cell fusion should therefore differ among taxa more widely than genes that control meiosis or vegetative functions. This is certainly true for phenotypes. Meiotic processes are rather uniform among eukaryotes, involving a stereotyped succession of chromosome movements. There are some exceptions to this generalization, in particular those taxa in which an achiasmate meiosis is accomplished in a single division, and Raikov (1982) has suggested that the diversity of meiotic processes alone indicates that sex has arisen several times among eukaryotes. Nevertheless, meiosis is strikingly similar in taxa as distant as veasts, angiosperms, and vertebrates. Mating systems, on the other hand, are extraordinarily variable, with different taxa having fundamentally different ways of bringing sexually competent cells into contact. Whether or not this diversity extends to mating-type sequences has not vet been carefully investigated, but it is a strong prediction of the theory being advanced here that it does. Similarities between the mating-type genes of different fungi have been detected. Within a genus, or set of related genera, either mating-type idiomorph is generally very similar in different species (see Metzenberg and Glass 1990) and has no doubt descended from a common ancestral sequence. The situation in more distantly related taxa is less clear. One mating type of Podospora could be cloned because it hybridized with a probe containing the A mating-type sequence of Neurospora, and it is possible that the two are homologous. On the other hand, a probe containing the a matingtype sequence of Neurospora hybridized with sequences present in both mating types of *Podospora*. This suggests that the second Podospora mating type has been acquired independently (Picard et al. 1991) and perhaps that the Neurospora a gene became fixed in *Podospora* at some time in the past, and now represents a fossil of that ancient infection. The *a* gene product also has about 30% similarity (of amino acids) over part of its length to the entire product of the MAT-M gene of Schizosaccharomyces pombe (Staben and Yanofsky 1990). On the other hand, the Neurospora A gene shares a short 15-amino acid region of similarity with the MAT α 1 gene of Saccharomyces cerevisiae (Glass et al. 1990). Whether or not these similarities indicate true homology, rather than common features of systems designed to facilitate horizontal transfer, is not known. However, Saccharomyces and Schizosaccharomyces are not closely related; if the Neurospora mating-type idiomorphs have really descended from ancestral sequences in these two yeasts, it is likely that they were acquired independently and by some form of horizontal transfer. The situation should be clarified in the near future by sequence information from chlorophytes, myxomycetes, and ciliates.

Mating-type genes in distantly related organisms are similar in one respect-that their expression is usually elicited by the stress associated with crowding or starvation. This is true among metazoans as well as among microbes (see Bell 1982) and has been interpreted to mean that it is advantageous for individuals (or, strictly, lineages) to diversify offspring at the end of one growing season in preparation for the unpredictable conditions of the next (Williams 1975). However, this pattern is also compatible with a parasite origin of mating-type genes. So long as conditions remain favorable and cells are growing vigorously, a parasite may maximize its fitness by remaining within the rapidly proliferating lineage. When conditions worsen, a parasite can detect the external change only as a change in the internal state of the cell, which will sometimes signal a deterioration specific to that cell, either because it has recently undergone a deleterious mutation or because its particular genotype is not well adapted to a change in external conditions to which most of the rest of the population is insensitive. It will then, on average, pay a parasitic sequence to transfer itself to a new lineage, which is likely to be unmutated or well adapted. It may, of course, pay to direct mating whenever the opportunity arises, for example if compatible mating partners are only rarely encountered, but to the extent that environmental variance in expression does evolve, nutritional stress is an appropriate cue for a parasite to use.

If mating-type genes arose as genome parasites, it might be possible to create new mating specificities by transforming cultures (e.g., a single genotype of a heterothallic form) or even perennially asexual populations (e.g., of *Euglena* or *Chlorella*) with cloned DNA representing the mating-type genes of an unrelated organism. Experiments of this sort are now technically feasible, though so far as I know none have yet been done. Unfortunately, failure would not be decisive, since even if integrated the cloned fragments might lack appropriate conditions for expression. However, if they succeeded they would provide very convincing circumstantial evidence for the plausibility of the theory.

Dynamics of Highly Multipolar Systems

Isogametic chlorophytes, most ascomycetes, and some basidiomycetes are bipolar with respect to mating, whether or not more complex compatibilities are required for subsequent development, and in all cases so far investigated the genes responsible are idiomorphic. In other organisms, such as most basidiomycetes, ciliates, and plasmodial slime-moulds, gender may be highly multipolar, with dozens or hundreds of different mating types. Idiomorphic mating-type genes in bipolar systems have been interpreted as descending from unrelated infectious elements. It does not seem very likely that organisms with multipolar mating systems have been infected on dozens of different occasions. At least some genes in multipolar systems may be idiomorphic, but others, as in Coprinus, are likely to be allelic. I expect that most differences between mating-type genes in multipolar mating systems will turn out to be allelic in nature, although there may be two or more groups of allelic genes, which represent radiations from idiomorphic ancestors. How do such allelic variants arise and spread?

It is a commonplace, though rarely quantified, observation that pure cultures which are maintained vegetatively tend to lose the ability to mate (for references, see Bell, 1992). Da Silva and Bell (1992) maintained sexual and asexual lines of Chlamydomonas under identical conditions, recruiting the sexual lines exclusively from zygotes every 10 or so doublings, while perpetuating the asexual lines without mating. When the asexual lines were eventually allowed to mate, after about 100 doublings, zygote production was less than 1% of that in the sexual lines. It seems likely, therefore, that sex is readily and rather rapidly lost in natural populations. It is possible for this to occur, because the sexual and vegetative cycles are distinct in eukaryotic microbes, rather than being conflated, as they are in most plants and animals, so that reproduction is not defective in sexually sterile lineages. Two processes may contribute to the loss of sexual competence. First, mutations of mating-type genes are likely to be neutral during strictly vegetative proliferation. They will therefore accumulate, with different lineages independently acquiring different mutations, yielding impaired phenotypes when mating is elicited and the mutated genes expressed. Second, because the sexual and vegetative cycles are separate, they can evolve independently, so that genetic changes which increase the rate of vegetative proliferation may be favored, even if they reduce mating ability. Da Silva and Bell (1992) found that vegetative rates of increase were lower in their sexual than in their vegetative lines, suggesting the operation of this sort of antagonistic pleiotropy. The outcome of mutation accumulation and antagonistic pleiotropy is that most extensive collections of sexual microbes include a substantial minority of sexually sterile clones, morphologically identical to their sexual relatives, but unable to mate with them. Moreover, it is usually found that when a rare mating does occur, subsequent development and meiosis proceed normally, suggesting that damage to meiotic systems is much less frequent or severe than damage to fusion systems. The occurrence of sexually sterile isolates is reviewed by Bell (1992).

Sexual sterility is presumably caused by mutations, perhaps involving cell-surface molecules involved in gametic recognition and interaction. In some cases these will unconditionally negate any possibility of sexual interaction, but in other cases complementary mutations in another lineage may restore mating ability by creating a novel sexual specificity. Complementary mutations of this sort will spread in a manner similar to parasitic elements. A sexually sterile lineage is not vegetatively impaired and may proliferate successfully and become very abundant. A mating-type gene, which, by mutation, acquires the ability to mate with members of this lineage, will automatically tend to increase in number; from the point of view of such a gene, the previously sterile lineage represents a large new cytoplasmic resource to be exploited. Such genes are not parasites, in the strict sense of originating as autonomously replicating genetic elements, but their dynamics will be very similar to those of strictly parasitic elements. If this speculation is true, then mutagenesis, which merely destroys sexual competence in bipolar idiomorphic forms, will occasionally create new mating specificities in highly multipolar taxa. Collins and Gong (1985) reported that one line of an apomictic strain of the myxomycete Didymium, in which there is a long series of homologous mating-type genes, spontaneously reverted to heterothallic sexuality, and this may represent an example of the acquisition of a novel sexual specificity. The interpretation that I have outlined should be contrasted with the more conventional view that multipolar systems benefit the individual by ensuring a high probability of wide outcrossing.

The frequent loss and restoration of sexual contact between lineages implies that the genetic population structure of eukaryotic microbes will be very different from that of large multicellular organisms. Rather than sexually isolated species each comprising two highly differentiated genders, the coupled processes of repeated isolation and reinvasion will produce much more diffuse assemblages of partially or asymmetrically isolated sets of lineages, many of which will be unable to mate with one another, but which nevertheless exchange genes through a series of intermediate matings. Such taxa are reticulate, forming a network whose connections continually change through time, rather than forming a tree of permanently isolated clades. I have documented elsewhere (Bell, 1992) the widespread and perhaps general occurrence of reticulate taxa among protists, where they stand as a witness to the complex and shifting patterns of gender in natural populations of eukaryotic microbes.

Parasite Origin of Meiotic Processes

Meiosis or some equivalent process is the inevitable concomitant of fusion in the sexual cycle. In most organisms, meiosis involves the orderly segregation of chromosomes, mediated by centromeres, and the recombination by crossing over of chromosomal genes. The entire process can be viewed as an intricate but harmonious device which benefits the genome, and thus the individual, as a whole, despite continuing uncertainty about the function of genetic recombination. However, it is also conceivable that it represents the stabilization of genetic elements which were originally parasites.

Centromeres

The molecular structure of centromeres has been worked out in yeasts (review in Schulman and Bloom 1991). In Saccharomyces the centromere is an extremely short (about 125 bp) unique sequence comprising an ordered series of three elements of different size: alterations in two of these somewhat reduce the fidelity of segregation, while alterations in the third, or changes in the arrangement of the three elements, are much more severely deleterious. In Schizosaccharomyces the centromere comprises a unique central region flanked by inverted repeats, the structure and arrangement of which are essential for normal meiotic behavior. The centromeric sequences of these two ascomycetes are thus completely different in structure.

Very short artificial chromosomes can be made by ligating a centromere, an origin of replication, and telomeres. When introduced into yeast cells there will replicate and segregate successfully, although the fidelity of segregation is often rather low; such yeast artificial chromosomes are now a familiar tool of molecular genetics. Imagine the behavior of an autonomous parasitic plasmid within a cell. If only a few copies are maintained per cell, the plasmid is likely to be lost by chance when the cell divides, with all copies being received by one daughter and none by the other. Since loss is much more likely than reinvasion, such plasmids will readily become extinct. If many copies are maintained, stochastic loss is much less likely, but such a large quantity of parasitic DNA imposes a burden on the cell that may reduce its rate of increase, and thus reduce the rate at which the plasmid itself proliferates. A plasmid that incorporates a centromere-like sequence will avoid both problems by avoiding stochastic loss while maintaining low copy number. There are, of course, many ways of stabilizing plasmid segregation, but centromeres, with their ability to nucleate microtubules, may provide the only one capable of moving massive quantities of DNA, and thus the only one that could evolve into the characteristic eukaryote mechanism. I suggest that centromeres originated as sequences which stabilized the segregation of low copy number plasmids, thereby enhancing their effectiveness as intracellular parasites. The brevity of Saccharomyces centromeres suggests that such structures may arise rather readily. Their ability to stabilize short artificial chromosomes shows that the scheme put forward here is not inherently implausible.

In most sexual organisms, meiosis is a complicated process involving two nuclear divisions and must have evolved from some simpler process. The relocation of nuclear genes to centromere-bearing plasmids would lead automatically to a crude, single-division reductional meiosis, given homologue pairing and the suppression of DNA synthesis before cytokinesis, both of which are easily acquired processes that would be very strongly favored by selection at the level of the cell. Single-division meioses are known from a number of phyla, including Dinoflagellata, Zoomastigina, and Sporozoa. With no DNA replication, there is no stage at which bivalents comprise four chromatids, and thus no crossing over. Meiosis therefore probably evolved as a simple ploidy-reducing device before it became associated with genetic recombination. I shall suggest that crossing over may have arisen when primitive meiotic organisms were infected by a new category of parasites capable of exploiting this new cytology.

Telomeres

The linear chromosomes of eukaryotes are made possible by specialized terminal structures, the telomeres. Telomeric DNA consists of simple repeated sequences, and thus resembles the repeated sequences which in many organisms occur abundantly throughout the genome. These are, for the most part, generated by the reverse transcriptases encoded by retroviruses or similar elements. It has recently been discovered that telomerase, the enzyme responsible for the production of telomeres, is a ribonucleoprotein whose RNA moiety includes a template for telomeric DNA; in other words, it is a specialized reverse transcriptase (Shippen-Lentz and Blackburn 1990). The mode of formation of telomeres makes it conceivable that they were originally acquired as sequences encoded by retroviruses or retroposons.

Crossing Over

The exchange of material between chromosomes by crossing over may stabilize meiotic bivalents and facilitate disjunction. Alternatively, crossing over leads to genetic recombination, which may evolve because selection favors the production of progeny which differ from their parents or among themselves. A third possibility was suggested by Bell and Burt (1990): genes that promote crossing over may be favored if they are able systematically to transfer themselves from less successful to more successful chromosomes. Consider a gene that directs crossing over, but is expressed only in sick or stressed cells. The argument is similar to that advanced in the case of genes which direct sexual fusion. The signal perceived by such a gene might indicate a general deterioration of the environment, but might also indicate a lesion specific to the cell, and perhaps to the chromosome on which the gene is located. It may therefore be advantageous on average for such a gene to move from its present location to the homologous location on the chromosome with which it is paired, a rat leaving a sinking ship. This represents a theory of inducible recombination which interprets at least the initial evolution of crossing over as a process of opportunistic infection, without regard to its consequences for the rest of the genome.

To the extent that this original property is retained by contemporary genes, the theory makes two explicit predictions about such genes. The first is that there should be environmental variance in recombination rates, which should be higher in more stressful environments. The second is that genes which direct crossing over should be located at the ends of chromosomes, near to the telomere, and should cause chiasmata to form as close as possible to their boundary on the centromere side, since they should be selected to carry with them as little material as possible from the chromosome they are leaving. Fungal genes that alter rates of recombination are not in general located near the ends of chromosomes, and they may act specifically to alter rates of recombination on other chromosomes (review in Catcheside 1977), two facts which are very unfavorable to a theory of inducible recombination. However, such genes are detected by mutations that reduce rates of recombination and may represent loss-of-function changes of little relevance to genes which are selected specifically for their ability to facilitate their own transfer via crossing over. The frequent and widespread occurrence of terminal and subterminal chiasmata, which can do little to recombine genes, is suggestive but inconclusive.

A similar interpretation of recombination was suggested by Hickey and Rose (1988). New copies of transposons often cause mutations when they reintegrate, thus creating selection against high rates of transposition at the level of the cell. Transposons that promote crossing over will tend to spread because the source copy is likely to be separated by recombination from the transposed copy, and then will not experience the cellular side effects of transposition. While it is known that transposons often increase rates of recombination, no unambiguous experimental test of this elegant idea has yet been reported.

Gene Structure

The protein-coding sequence of eukaryotic genes is often interrupted by introns. It is by now a familiar idea that introns may be the hulks of formerly active parasites, or may even still be capable of actively maintaining themselves (see Hickey and Benkel 1985). Elements such as the mitochondrial α sequence of *C. smithii* are now providing concrete evidence that introns should be interpreted as sexual parasites.

The Domestication of Parasitic Elements

Horizontally transmitted elements will tend to be more or less malignant, from the point of view of other genes in the same line of descent, to the extent that they evolve greater infective ability at the expense of reducing vegetative proliferation. Vertically transmitted elements will be benign, since they can evolve greater rates of increase only by enhancing the success of the lineage as a whole. In a series from bacteria through protists to multicellular eukaryotes, there is a tendency for genes involved in sexual behavior, originally spreading horizontally as parasites, to become progressively domesticated by vertical transmission.

In a population of primitive eukaryotes, elements directing horizontal transfer would act as simple parasites, infecting by transposition. Integrated into chromosomes, they could ensure their transmission only by directing the complete cell fusion characteristic of eukaryotes. Before the evolution of regular segregation, there would be a considerable risk that cell division would result in daughter cells with unbalanced or incomplete genomes. There would therefore be selection for uninfected cells to evolve resistance to fusion, and then of course for the genes directing fusion to overcome this resistance. The outcome of this coevolutionary struggle would be a great diversity of fusion systems, perhaps leading to the great diversity of modern mating systems. The conflicts of interest that arise when two or more elements infect the same population can also be resolved in different ways: by heterothallism, with different elements being partitioned to separate nuclei at meiosis; by homothallism, with both elements simultaneously present and expressed; or by switching, with both elements present but only one expressed.

Genes that direct meiosis are vertically transmitted, in that, when fusion has occurred, it is in the interests of the genome as a whole that homologous elements are assorted regularly. Once centromere-like sequences have evolved as devices to stabilize the distribution of parasitic plasmids among daughter cells, other genes will benefit by being linked to them, building up the first eukaryotic chromosomes. This might occur very readily, in view of the fact that larger yeast artificial chromosomes tend to segregate more reliably than shorter segments, perhaps simply because the homologues are paired for longer. There is no conflict of interest, but rather the reverse, between the genome as a whole and elements that direct proper recovery from fusion, and the mutualistic relationship between the two will lead to a uniformity of meiotic behavior. The sexual cycle of eukaryotes, according to this view, begins to assume its modern form when the deleterious effects of one genomic parasite are cured by the intervention of a second. Whether primitive centromeres participated in a mitotic process during vegetative proliferation, before being co-opted into the resolution of genomes in fused cells, is difficult to say; no doubt meiosis and mitosis were evolving at the same time, being organized to a large extent by the same set of genes.

The activity of mating-type genes may be antagonistic to the interests of the rest of the genome to the extent that sex and reproduction are decoupled, as they are in most protists. In large multicellular organisms, on the other hand, sex and reproduction are obligately associated. The parasitic nature of mating-type genes is then finally lost, because, being transmitted in the same way, their interests no longer conflict with those of other genes, and they behave as part of an intricate, harmoniously functioning whole. Sex in such organisms must be maintained, and the spread of purely asexual variants suppressed, because its side effects benefit the genome as a whole. It is with such processes that most of the papers at this meeting have been concerned. It would be pleasingly symmetrical if the sexual cycle of eukaryotes, having evolved by the domestication of a zoo of genomic parasites, now functions itself to check the ravages of worms and mould.

References

Baldauf SL and Palmer JD, 1990. Evolutionary transfer of the chloroplast *tufA* gene to the nucleus. Nature 344: 262–265.

Bannister JV and Parker MW, 1985. The presence of a copper/zinc superoxide dismutase in the bacterium *Photobacter leignathi*: a likely case of gene transfer from eukaryotes to prokaryotes. Proc Natl Acad Sci USA 82: 149–152.

Beach DH and Klar AJS, 1984. Rearrangements of the transposable mating-type cassettes of fission yeast. EMBO J 3:603-610.

Bell G, 1982. The masterpiece of nature. London: Croom Helm and Berkeley: University of California Press.

Bell G, 1992. The emergence of gender and the nature of species in eukaryotic microbes. Verh Dtsch Zool Ges 85:161-175.

Bell G and Burt A, 1990. B-chromosomes: germ-line parasites which induce changes in host recombination. Parasitology 100:519-526.

Bell G and Koufopanou V, 1991. The architecture of the life cycle in small organisms. Phil Trans R Soc Lond B 332:81–89.

Benveniste RE, 1985. The contribution of retroviruses to the study of mammalian evolution. In: Molecular evolutionary genetics (McIntyre RJ, ed). New York: Plenum; 95-136.

Bölker M, Urban M, and Kahmann R, 1992. The *a* mating-type locus of *U. maydis* specifies cell-signalling components. Cell 68:441-450.

Boynton JE, Harris EH, Burkhart BD, Lamerson PM, and Gillham NW, 1987. Transmission of mitochondrial and chloroplast genomes in crosses of *Chlamydomonas*. Proc Natl Acad Sci USA 84:2391-2395.

Carlson TA and Chelm BK, 1986. Apparent eukaryotic origin of glutamine synthetase II from the bacterium *Bradyrhizobium japonicum*. Nature 322:568-570.

Catcheside DG, 1977. The genetics of recombination. London: Edward Arnold.

Charlesworth B and Langley CH, 1986. The evolution of self-regulated transposition of transposable elements. Genetics 112:359-383.

Clewell DB and Gawron-Burke C, 1986. Conjugative transposons and the dissemination of antibiotic resistance in streptococci. Annu Rev Microbiol 40:635-659.

Colleaux L, Michel-Wolwertz M-R, Matagne RF, and Dujon B, 1990. The apocytochrome *b* gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. Mol Gen Genet 223:288–296.

Collins OR and Gong T, 1985. Genetical relatedness of a former apomict and a heterothallic isolate in *Didymium iridis* (Myxomycetes). Mycologia 77:300-307.

Cornish-Borden A, 1985. The amino-acid sequences of the copper/zinc superoxide dismutases from swordfish and *Photobacter leignathi* confirm the predictions made from their compositions. Eur J Biochem 151:333-335.

Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, and Chovnick A, 1990. Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species. Genetics 124:339–355.

Da Silva J and Bell G, 1992. The ecology and genetics of fitness in *Chlamydomonas*. VI. Antagonism of sexual and natural selection. Proc R Soc Lond B 249:227-233.

Doolittle RF, Feng DF, Anderson KL, and Alberro MR, 1990. A naturally-occurring horizontal gene transfer from a eukaryote to a prokaryote. J Mol Evol 31:383-388.

Egel R, Nielson O, and Weilguny D, 1990. Sexual differentiation in fission yeast. Trends Genet 6:369-373.

Farrelly F and Butow RA, 1983. Rearranged mitochondrial genes in the yeast nuclear genome. Nature 301: 296-301. Froeliger EH and Leong SA, 1991. The *a* mating-type alleles of *Ustilago maydis* are idiomorphs. Gene 100: 113–122.

Giasson L, Specht CA, Milgrim C, Novotny CP, and Ullrich RC, 1989. Cloning and comparison of A matingtype alleles in the basidiomycete Schizophyllum commune. Mol Gen Genet 218:72-77.

Glass NL, Grotelueschen J, and Metzenberg RL, 1990. *Neurospora crassa A* mating-type region. Proc Natl Acad Sci USA 87:4912-4916.

Heim L and Gutz H, 1991. Rescue of lethal mutations in the mating-type region of *Schizosaccharomyces pombe* by an extrachromosomal element. Curr Genet 19:435-444.

Heinemann JA and Sprague GFJ, 1989. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. Nature 340:205-209.

Herskowitz I, 1988. Life cycle of the budding yeast Saccharomyces cerevisiae. Microbiol Rev 52:536-553.

Hickey DA, 1982. Selfish DNA: a sexually-transmitted nuclear parasite. Genetics 101:519-531.

Hickey DA and Benkel BF, 1985. Splicing and the evolution of introns. Science 316:582.

Hickey DA and Rose MR, 1988. The role of gene transfer in the evolution of eukaryotic sex. In: The evolution of sex (Michod RE and Levin BR, eds). Sunderland, Massachusetts: Sinauer; 161–175.

Jacquier A and Dujon B, 1985. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell 41:383–394.

Kawano S, Takano H, Mori K, and Kuroiwa T, 1991. A mitochondrial plasmid that promotes mitochondrial fusion in *Physarum polycephalum*. Protoplasma 160:167-169.

Kelly M, Burke J, Smith M, Klar A, and Beach D, 1988. Four mating-type genes control sexual differentiation in the fission yeast. EMBO J 7:1537–1547.

Kronstad JW and Leong SA, 1990. The *b* mating-type locus of *Ustilago maydis* contains variable and constant regions. Genes & Dev 4:1384–1395.

Leunissen JAM and de Jong WW, 1986. Copper/zinc superoxide dismutase: how likely is gene transfer from ponyfish to *Photobacterium leignathi*? J Mol Evol 23: 250-258.

Margulis L, 1981. Symbiosis in cell evolution. San Francisco: W. H. Freeman.

Marshall P and Lemieux C, 1991. Cleavage pattern of the homing endonuclease encoded by the fifth intron in the chloroplast large subunit rRNA-encoding gene of *Chlamydomonas eugametos*. Gene 104:241–245.

Martin JP and Fridovich I, 1981. Evidence for a natural gene transfer from the ponyfish to its bioluminescent bacterial symbiont *Photobacter leiognathi*. J Biol Chem 256:6080–6089.

May G, Le Chavanton L, and Pukkila PJ, 1991. Molecular analysis of the *Coprinus cinereus* mating type A factor demonstrates an unexpectedly complex structure. Genetics 128:529–538.

Maynard Smith J and Nee S, 1989. The evolutionary biology of molecular parasites. Symp Br Soc Parasitol 78–92.

Metzenberg RL, 1990. The role of similarity and difference in fungal mating. Genetics 125:457-462.

Metzenberg RL and Glass NL, 1990. Mating type and mating strategies in *Neurospora*. BioEssays 12:53-59.

Mizrokhi LJ and Mazo AM, 1990. Evidence for horizontal transmission of the mobile element *jockey* between distant *Drosophila* species. Proc Natl Acad Sci USA 87:9216-9220.

Mutasa ES, Tymon AM, Gottgens B, Mellon FM, Little PFR and Casselton LA, 1990. Molecular organisation of an *A* mating type factor in the basidiomycete fungus *Coprinus cinereus*. Curr Genet 18:223–229.

Nagley P and Devenish RJ, 1989. Leading organellar proteins along new pathways: the relocation of mitochondrial and chloroplast genes to the nucleus. Trends Biochem Sci 14:31-35.

Nishikawa M, Suzuki K, and Yoshida K, 1992. DNA integration into recipient yeast chromosomes by transkingdom conjugation between *Escherichia coli* and *Saccharomyces cerevisiae*. Curr Genet 21:101-108.

Picard M, Debuchy R, and Coppin E, 1991. Cloning the mating types of the heterothallic fungus *Podospora anserina*: developmental features of haploid transformants carrying both mating types. Genetics 128:539-547.

Raikov IB, 1982. The protozoan nucleus. Trans. Bobrov N and Verkhovtseva M. Vienna: Springer.

Remacle C, Bovie C, Michel-Wolwertz M-R, Loppes R, and Matagne RF, 1990. Mitochondrial genome transmission in *Chlamydomonas* diploids obtained by sexual crosses and artificial fusions: role of the mating type and of a 1 kb intron. Mol Gen Genet 223:180-184.

Rose MR, 1983. The contagion mechanism for the evolution of sex. J Theoret Biol 101:137-146.

Schulman I and Bloom KS, 1991. Centromeres: an integrated protein/DNA complex required for chromosome movement. Annu Rev Cell Biol 7:311-336.

Shatters RG and Kahn ML, 1989. Glutamine synthetase II in *Rhizobium*: re-examination of the proposed horizontal transfer of DNA from eukaryotes to prokaryotes. J Mol Evol 29:422–428.

Shippen-Lentz D and Blackburn EH, 1990. Functional evidence for an RNA template in telomerase. Science 247:546–552.

Staben C and Yanofsky C, 1990. *Neurospora crassa a* mating-type region. Proc Nat Acad Sci USA 87:4917-4921.

Stern DB and Palmer JD, 1984. Extensive and widespread homologies between mitochondrial DNA and chloroplast DNA in plants. Proc Natl Acad Sci USA 81: 1946-1951.

Stotzky G and Babich H, 1986. Survival of, and genetic transfer by, genetically engineered bacteria in natural environments. Adv Appl Microbiol 31:93–138.

Trieu-Cuot P, Carlier C, and Courvalin P, 1988. Conjugative plasmid transfer from *Enterococcus faecalis* to *Escherichia coli*. J Bacteriol 170:4388–4391.

Wakabayashi S, Matsubara H, and Webster DA, 1986. Primary sequence of a dimeric bacterial haemoglobin from *Vitreoscilla*. Nature 322:481-483.

Williams GC, 1975. Sex and evolution. Princeton, New Jersey: Princeton University Press.

Zambryski P, 1988. Basic processes underlying *Agrobacterium*-mediated DNA transfer to plant cells. Annu Rev Genet 22:1-30.