Meiosis and the Evolution of Recombination at Low Mutation Rates

Damian D. G. Gessler and Shizhong Xu

Department of Botany and Plant Sciences, University of California, Riverside, California 92521

Manuscript received September 1, 1999
Accepted for publication May 26, 2000

ABSTRACT

The classical understanding of recombination is that in large asexual populations with multiplicative fitness, linkage disequilibrium is negligible, and thus there is no selective agent driving an allele for recombination. This has led researchers to recognize the importance of synergistic epistatic selection in generating negative linkage disequilibrium that thereby renders an advantage to recombination. Yet data on such selection is equivocal, and various works have shown that synergistic epistasis per se, when left unquantified in its magnitude or operation, is not sufficient to drive the evolution of recombination. Here we show that neither it, nor any mechanism generating negative linkage disequilibrium among fitness-related loci, is necessary. We demonstrate that a neutral gene for recombination can increase in frequency in a large population under a low mutation rate and strict multiplicative fitness. We work in a parameter range where individuals have, on average, less than one mutation each, yet recombination can still evolve. We demonstrate this in two ways: first, by examining the consequences of recombination correlated with misrepaired DNA damage and, second, by increasing the probability of recombination with declining fitness. Interestingly, the allele spreads without repairing even a single DNA mutation.

The evolution of recombination has historically been dominated by two main understandings. The first is that recombination destroys linkage disequilibrium; thus, if there is no linkage disequilibrium among fitness-related loci, there is no mechanism other than drift driving an allele for recombination. The second is that concurrent with the destruction of such linkage disequilibrium, there is a commensurate change in population mean fitness. A rise in mean fitness is often inferred, though is not sufficient, to conclude an adaptive advantage of recombination. These properties are based on well-established and well-verifed analyses (e.g., Fisher 1930; Felsenstein 1965; Maynard-Smith 1968; Kondrashov 1982; Charlesworth 1990; Barton 1995), and there is no question as to the truthfulness of the derivations as based on the assumptions of the underlying models. Importantly, these analyses have given us a fundamental understanding of the population genetic consequences of recombination. Yet one can ask, What is the dependency of the conclusions to those underlying assumptions? Specifically, consider recombination to be differentially induced (instead of being solely a function of the distance between two loci), for example, induced differentially as a function of a particular cellular state. To see how this question can be relevant for the evolution of meiosis, we start with the difference in neutral fixation probabilities between asexual and sexual populations. In this article, when we consider recombination, we are primarily interested in the spread of reciprocal, homologous recombination (as is exemplary of meiosis), though much of the model is feasibly applicable to prokaryotic recombination as well.

To begin, consider that in any population heritable differences in fitness mean that some individuals leave more offspring than others and that this bias is transmitted with some degree of error from generation to generation. This means that without recombination, the probability of fixation for all loci on the best haplotypes is disproportionately biased toward unity (Fisher 1930; Figure 1). This phenomenon is called background trapping. With recombination, genomes become mixed, and an allele’s fate becomes largely independent of its original genomic background (Table 1).

Inspection of Table 1 shows that asexual populations create an inherent selective pressure on genes in inferior haplotypes to escape their genetic background. Only among the best individuals is it advantageous for loci to remain linked. This is a direct and inescapable consequence of inheritance in asexual populations. To the degree that the best individuals constitute a subclass of the population (even if not necessarily few in number), this selection can be strong, since virtually all haplotypes save the best are destined for extinction (Fisher 1930; Williams 1975; Peck 1994). Yet despite this, escaping one’s genetic background is beyond the capability of most loci, for they code for no such enzyme to induce the behavior. But for genes that induce recombination, this is exactly what they do: they are themselves the genes that orchestrate mixis. Additionally, one would expect other genes to evolve cooperation with
TABLE 1

Probabilities of fixation for a neutral allele as a function of the individual within which it arose

<table>
<thead>
<tr>
<th>Ancestral individual</th>
<th>Mean probability of fixation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without recombination</td>
<td>1 ( \geq u_i )</td>
</tr>
<tr>
<td>With recombination</td>
<td>( \sim 1/N )</td>
</tr>
</tbody>
</table>

Ancestral individuals are ranked in terms of fitness from best (1) to worst (N), where \( u_i \) is the probability of fixation for an allele conditional on it arising in the \( i \)-th ranked ancestral individual. Probabilities are for haploid populations of size \( N \). The table reflects the fact that while the mean probability of fixation for neutral alleles is independent of linkage (Birky and Walsh 1988), the variance is not. Individuals of similar fitness may be grouped together into a class, the highest fit of which is called the “best class.” Classes can be in the thousands of individuals while still maintaining a bias in the distribution of probabilities between classes. In the simulations reported here, the best class is \( \sim 58,000 \) individuals.

These genes for recombination, since all genes on inferior haplotypes benefit from mixis. Note that this is entirely independent of, and in addition to, the classical arguments based on the law of large numbers that show an absence of linkage disequilibrium in large asexual populations. In fact, it works best in large populations, since it is the manifestation of the deterministic expectation that the best individuals monopolize fixation events that skews the probabilities in Table 1.

In natural systems, the idealized probabilities in Table 1 will be obfuscated by recurrent mutation, compensatory and beneficial mutations, and environmental changes. Yet since nonrecombining genomes have dynamics analogous to single-locus models with temporally varying selection coefficients, and since such models predict that haplotypes deterministically differ in their fixation probabilities so long as differences in their geometric mean fitness exceed \( N e^{-1} \), it is inescapable that asexual populations—at least as we currently understand them—will incorporate some bias in their distribution of fixation probabilities. The question then becomes, Is there sufficient genetic variance for alleles to exploit this variance in fixation probabilities to increase their own probabilities of fixation? Clearly, if inferior haplotypes are destined to extinction, then it is advantageous for their alleles to destroy their current association.

**MUTATION AND RECOMBINATION**

Currently, and with the exception of examples such as Weigle reactivation [i.e., SOS-induced viral reactivation via recombinatoric pathways (Weigle 1953; e.g., Calsou and Salles 1991 and references therein)], we know of no direct evidence that unicellular haplotypes differentially recombine based on their expected fitness. For ancient unicellulars, a simple mechanism would be for them to recombine if they acquired new mutations, since statistically this would coincide with their leaving the best class. While the hypothesis does not require that such a mechanistic correlation exists, here we review evidence that it may.

There is strong evidence that ionizing radiation, mutagens, and cellular processes themselves cause double-strand breaks (DSBs); these act as initiation sites for recombination (Resnick 1976; Thaler and Stahl 1988; Lichten and Goldman 1995; Shinohara and Ogawa 1995; Stahl 1996; Nicolas 1998). There is then empirical evidence that DNA damage specifically induces recombination in modern-day organisms. Among the multiple cellular responses to DNA damage (Thacker 1999), recombination is a specific response to both induced (Eckardt-Schupp and Klaus 1999) and spontaneous damage (Sonoda et al. 1998), with DSB initiation of recombination being observed across kingdoms (Cox 1997; Dernburg et al. 1998). Additionally, single-stranded lesions may also invoke recombinatoric repair (Raderschall et al. 1999). Genes involved in recombination induction, such as mammalian homologues of the yeast Rad51, are upregulated after DNA damage (Haaf et al. 1995) and, upon experimental over-
expression, confer elevated levels of recombination and enhanced radiation resistance (Vispe et al. 1998). Repair is generally homologous in yeast and sometimes reciprocal (Liang et al. 1998), homologous reciprocal recombination having evolved earlier in prokaryotes to repair DSBs created during replication (Friedberg et al. 1995). This is an important point. Thus, while there are certainly nonrecombinogenic repair pathways (Kanaar et al. 1998; Thacker 1999), there is a direct link between DSBs and wild-type cells actively invoking homologous recombination upon detection (Kanaar et al. 1998).

Because the correction of DSBs itself can be mutagenic (Strathern et al. 1995; Holbeck and Strathern 1997; also Kuzminov 1999), cells that experience more lesions stand to experience more incorporated mutations and more recombination. The link between mutation and recombination is correlative, but the link between DNA lesions and either is causative.

In addition to the link between DNA damage and recombination, it would be pertinent to know if damage increased the likelihood of recombination even among undamaged sites. Again, there is supportive evidence: Golub and Low (1983), working with Escherichia coli, describe how the “interaction of a damaged DNA molecule with a homologous undamaged one . . . makes the latter a more active recombination substrate” with “the induction of recombination events at sites far removed from the sites of interaction” (p. 1405). Similar evidence is found in eukaryotes. Fabre and Roman (1977) crossed irradiated haploid yeast with nonirradiated diploid yeast to produce triploid offspring. By using genetic markers, they noted a substantial increase in recombination among the nonirradiated diploid pair of chromosomes in the triploid zygotes. This experiment was specifically designed to separate the conditions of damage-inducing recombinatoric repair at the lesion vs. inducible recombinatoric competency for the cell itself. It is the latter that is of particular interest here. When Fabre and Roman allowed the haploids to correct newly created pyrimidine dimers via photoreactivation (after irradiation but before mating), they noted a corresponding reduction in recombination among the nonirradiated diploid pair. Interestingly, when they repeated the experiment using kar− mutants that inhibited the fusion of the diploid and haploid nuclei in the triploid heterokaryon, they still observed elevated levels of recombination in the nonirradiated diploid nucleus.

The preceding studies do not show that inferior haplotypes are more likely to recombine. But they do suggest that biases in recombinatoric propensity exist: when we consider recombination theoretically, we should consider its differential induction. Currently, recombination within a cell is often viewed in terms of DNA repair—that is, for the benefit of the cell (though see Thaler 1994)—but it is neither predicated nor necessary that recombination between ancient unicellular haploids was so driven. Between unicellulars, recombination destroys haplotypes as linkage groups, so genes benefit from a coordinated induction whenever they stand to benefit from the destruction of their current association. From the perspective of linked genes, the detection of DNA damage is an important cue to allow them to avoid eventual extinction: either fix the DNA or leave the haplotype.

THE MODEL

To examine the consequences of the latter for the evolution of meiosis, we use a model of early eukaryotic, proto-meiotic, haploid organisms. We introduce genetic variance for recombination in the form of an invading allele that instigates homologous reciprocal recombination between unicellulars upon coupling. But the allele, instead of inducing recombination arbitrarily and being oblivious to the state of the cell within which it resides, is integrated into its cellular mechanisms: it is more likely to be activated in cells that have recently acquired deleterious mutations. We then relax this mutation/recombination link in a more general model that negatively correlates recombination with fitness.

In a computer, we construct a population of \( N = 10^5 \) asexual individuals. The computer keeps track of each mutation’s position and deleterious effect on a single haploid chromosome for each individual. Actual positions approximate an infinite-sites model. Each generation, individuals go through a stochastic process of mutation, selection, and reproduction. Mutations occur randomly throughout the genome with a Poisson mean of \( \mu = 0.0034 \) individual\(^{-1} \) generation\(^{-1} \) (Drake et al. 1998). Incoming deleterious selection coefficients are random variates from a negative exponential distribution of mean \( s = 0.02 \) (Ohta 1977; Gillespie 1991; Lynch et al. 1995, 1999). Because it is convenient to work with a distribution that has a unit integral on the support \([0,1]\), the negative exponential distribution is computationally approximated by using a beta distribution with shape parameters \( \alpha = 1 \) and \( \beta = s^{-1} - 1 = 49 \). There is currently much discussion in the literature about appropriate values and distributions of \( \mu \) and \( s \), but the important point in this parameterization is that both \( N\mu \) and \( N^3 s \) exceed unity (where \( s \) is the average segregating selection coefficient in the population), while \( \mu \) is in the range believable for DNA microbes. We do not explicitly model DSBs, in large part because we have little or no data on how their resolution generates a mutational spectrum of selection coefficients. We do know, though, that cells experience thousands of lesions per generation (Cox 1997; Gupta and Lutz 1999; Kuzminov 1999)—orders of magnitude higher than their net deleterious mutation rate—so we correlate recombination with this lower rate and then weaken this correlation further in later simulations.

Selection in the model is strictly multiplicative, with individuals that pass selection becoming the adults of the next generation. Before invading alleles are introduced, each population first comes to approximate mu-
tation-selection balance such that mean fitness equals $e^{-p}$ (generation 0 on graph). Variance for recombination is introduced by mutating an allele at a randomly chosen neutral locus, such that individuals with this allele may instigate coupling with another randomly chosen individual. Coupling is defined loosely, meaning two units unite, recombine their genomes, and disassociate. For the first set of runs (mutational-induction runs), an individual instigates recombination only if it both has the allele and has acquired a mutation in the current generation. Since the probability of a mutation is $1 - e^{-0.0034} \approx 0.0054$, this is a rare event. For the second set of runs (fitness-conditional runs), for each generation, each individual’s fitness is compared to a $0\sim1$ uniform variate. If its fitness is less than the variate and it has the recombination allele, then it instigates recombination. Since mean fitness is high ($e^{-0.0034} \approx 0.9966$), this also is a rare event. In contrast to the mutational-induction runs, induction of recombination is independent of mutational events per se and thus can be instigated by members of the best class. In additional runs, noise is added to reduce the correlation between fitness and recombination. Each individual’s fitness is mapped to an indicator variable (a 0/1 random variate) such that the correlation coefficient between it and an individual’s fitness is $\sim 0.01$. Individuals need both the correct state of the variate and the allele to instigate recombination.

Because of the low mean number of mutations per individual ($\mu/3 \approx 0.55$), recombination is modeled as free recombination. We did additional simulations with the number of chiasma as a Poisson random variate. As expected, one could demonstrate a decrease in the strength of selection on driving the recombination allele to fixation as the mean number of chiasma ($\eta$) approached 0 (for $\eta < 1$, data not shown), but the qualitative results are the same as those reported here.

In all treatments, the invading allele acts dominantly such that only one individual of a conjugating pair needs the allele for recombination to proceed. For controls, the computer also monitors an invading neutral allele. We ran three controls: (1) an invading recombination allele (but recombination is unconditional); (2) a neutral allele in a strictly asexual population; and (3) a neutral allele in a population already fixed for recombination. To most efficiently measure the realized strength of selection on the invading allele, the allele is introduced at a frequency of 0.5 and monitored for 1000 generations.

**RESULTS**

Figure 2 shows that in all treatment cases, the invading recombination allele sweeps rapidly toward fixation, converting the populations from asexual to sexual. The strength of selection is estimated from the deterministic prediction that $f_{1.5}(dx/sx(1-x)) = 1000$ generations implies $s = (\ln(p_0) - \ln(1 - p_0))/1000$, where $p_0$ is the final frequency. Accordingly, the estimated strengths of selection are $1.2 \times 10^{-3} \pm 2.8 \times 10^{-4}$ SD ($P = 1.5 \times 10^{-4}$), $1.1 \times 10^{-3} \pm 3.3 \times 10^{-4}$ SD ($P = 1.3 \times 10^{-5}$), and $3.3 \times 10^{-4} \pm 1.6 \times 10^{-4}$ SD ($P = 1.1 \times 10^{-3}$) for the mutational induction, fitness correlation, and fitness correlation with noise runs, respectively. $P$ values are two-tailed Behrens-Fisher probabilities for deviations from 0. No control runs differed significantly from 0 or each other.

When we modify the mutation-induction runs such that recombination is randomly induced on only 10% of mutations entering at the net deleterious rate, the allele still spreads. [After 10,000 generations, 10 out of 10 runs had frequencies $>0.5$ with 1 run reaching fixation. Of the 9 still segregating, mean frequency was $0.780 \pm 0.136$ SD ($P = 2.5 \times 10^{-4}$); not shown in Figure 2.]

Since there is no linkage disequilibrium in the populations before the introduction of the alleles, there is no change in the standard population-wide measures and there is no significant difference in either mean fitness or the coefficient of variation in fitness between any combination of treatments and controls ($P > 0.05$). We did additional simulations with alleles at high frequency to verify that mean fitness remained time invariant.

**DISCUSSION**

The allele modeled above purposefully does not “repair” mutations, and thus its advantage is not explained by an alteration of the segregating load; it merely recom-
bines out of inferior haplotypes as they tend to become mutated. This makes its mechanism of spread distinct from, yet not antagonistic to, traditional repair/recombination hypotheses (Szostak et al. 1983; Bernstein et al. 1985). All the same, the allele sweeps rapidly toward fixation; it does so in large populations at low mutation rates with neither epistasis nor a rise in mean fitness—the very parameter space long held immune to invasion by recombination.

We modeled these simulations in terms of the induction of recombination, yet the process can be equivalently viewed as the suppression of recombination by the best haplotypes. All genes on the best haplotypes—including the recombination allele itself—benefit by a coordinated induction/suppression rule as this is in the best interest of all linked alleles. Redfield (1988) examined a situation under the context of nonreciprocal recombination using a model of bacterial transformation. In that model, mean fitness increased and recombination spread because the best class both suppressed and inhibited transformation. Here, though, reciprocal recombination can always be forced upon the best class, so linkage disequilibrium is not generated in the same manner. It is plausible that if we allowed the best class to evolve recombination-resistant genes, recombination may not spread. But Redfield’s work shows that the fate of a recombination allele cannot be inferred without assumptions on its underlying genetics. With pleiotropic inhibition in the best class, the allele not only spreads but increases population mean fitness.

The simulations that negatively correlate recombination and fitness show that the mutation-recombination link is only a special case of a more general relationship. For example, we could have modeled a basal probability of recombination for all cells, with more mutated cells more likely to recombine. Any factor sufficiently increasing recombination with decreasing fitness will bias the probability of fixation for a background-sensitive recombination allele. The numerically small variance in fitness in these simulations ($\alpha^2 = e^{2\mu} [e^{\mu} - 1] = 6.75 \times 10^{-5}$) means that conditional induction and suppression cues can capitalize on subtle differences in fitness to produce a marked qualitative result.

Evolutionarily, the reliance on induction cues, while important, is a weak requirement. For recombination to spread, the haplotype need only (imperfectly) decide if it is likely to be the best or not the best. We assume this is never available per se, and thus the mutation-induction model is a hypothesis of how haplotypes could achieve this independent of any direct knowledge. But as $\mu$ increases, it becomes exponentially more likely that haplotypes will make the right decision by recombining regardless of the state of their genomes. The relationship is exponential because when stable, the relative size of the best class is $e^{\mu}$ ($1 - e^{-\mu}$)$^{-1}$ (Haigh 1978). Furthermore, any advantage to suppressing recombination drops to zero once $Ne^{-\mu/2}$ falls below unity (Gessler 1995). In this case, the best class itself is not stable, and even unconditional recombination is advantageous (Gessler and Xu 1999). This means that as $\mu$ increases, selection for conditional recombination will eventually exceed that of selection for unconditional isolation by the best class.

In all of the above cases, the allele spreads because it alters its probability of fixation (Table 1) by biasing the alteration of its genetic background. This generates positive linkage disequilibrium between the allele and its background, though not among fitness-related loci. At first, this absence of linkage disequilibrium among fitness-related loci may seem somewhat disquieting when extended as an explanation for the ubiquity of recombination. Our best evidence so far is that haploid genome-wide mutation rates of single-celled eukaryotes are on the order of 0.0034 (Drake et al. 1998). At rates just two- to fourfold higher, such as those possibly invoked by diplody, the mitigation of Muller’s ratchet can render a stable advantage to recombination even in the absence of epistasis (Gessler and Xu 1999). Thus there is only a small window between where recombination could have evolved due to the effects of background trapping and where higher mutation rates can maintain it because of its ability to destroy newly forming negative linkage disequilibrium. It is still unclear whether conditions favorable for gene assortment (such as synergistic epistasis) were always present yet unexploited before the evolution of recombination, or if conditions favorable for recombination arose along with the mutation rate.

That gene assortment is actively maintained is supported by molecular examinations in yeast. The yeast Saccharomyces cerevisiae create DSBs before meiosis in a strikingly nonrandom manner: 89% of DSBs on chromosome III are intergenic, and these DSBs correspond well with sites of crossing over (De Massy et al. 1994; Wu and Lighten 1994; Baudat and Nicolas 1997). Thus, for ancient unicellulars, even if recombination originally evolved for internal repair and spread due to background trapping, the subsequent regulation of self-induced DSBs could recruit it as an assortment process.

The fact that unrepaired DSBs are unequivocally deleterious makes it unlikely that cells would have created them without already having a robust mechanism to repair them in place. Thus we must hypothesize that DSB-induced recombination was already functional before we can attribute to recombination any benefits of assortment. This chronology is consistent with the demonstration of Gilbertson and Stahl (1994) that DSBs occur in haploid meiosis and do so preceding homologue pairing. Because intragenic crossing over is more likely to disrupt gene function than intergenic crossing over, an intergenic distribution of DSBs could evolve even under conditions of linkage equilibrium. (Clearly, coordinated attempts to dissolve linkage arrangements do little good if newly created haplotypes are even more heavily mutated.) Once $\mu$ rises such that $Ne^{-\mu/2} < 1$, the subsequent generation of negative linkage disequilib-
Figure 3.—Meiosis from the perspective of a gene for recombination. Cells in the top and bottom left are two organisms in the population. Asexually reproducing cells follow the top and bottom; the recombination pathway constitutes the middle. All three use the existing mitotic machinery to perform fission. Eventually, with the evolution of multicellularity (and possibly a sequestered germline), the dominant life phase may shift to the diploid stage marked “syngamy,” with the top- and bottom-extreme-left cells becoming a sperm and egg.

The evolution of meiosis: Figure 3 shows how the spread of a recombination allele as modeled in these simulations immediately generates a model of meiosis; i.e., the conversion of intracellular recombination to intercellular recombination can be considered the creative step in meiotic evolution (see also Maynard-Smith 1978, p. 8, and Ruvinsky 1997). Despite its simplicity, Figure 3 is not a null expectation, since knowledgeable authors have hypothesized the exact opposite: that meiosis I evolved before meiosis II.

Part of the difficulty in understanding meiosis has been that hypotheses have had to first advocate advantages to diploidy and have then needed to explain the subsequent and repetitive return of cells to haploidy (Maynard-Smith 1978, p. 9). The model presented here ignores the fitnesses of cells per se and concentrates solely on the selective conditions advantageous to the spread of a background-sensitive recombination allele. Thus, there is no conceptual problem with the coexistence of diploid and haploid states. The emphasis here is from the perspective of the fitness of the recombination allele (and all other genes that concomitantly benefit), with little regard for the effect of recombination on population mean fitness. It is thus neither dependent on, nor exclusionary to, hypotheses that predict that mean fitness will increase with transient diploidy, nor does it require any selective difference between the diploid and haploid states.

Figure 3 addresses two simple yet common questions: (1) In meiosis, why does the cell double its DNA only to cut it in half twice again, and (2) how could a process as complicated as meiosis—one involving dozens to hun-
dreds of genes—ever evolve? The hypothesis implied here is that meiosis has only become a gametogenic process, but it did not evolve that way. The duplication of DNA is inherently mutagenic, as is the correction of lesions prior to replication (Kuzminov 1999). This places selection on DNA repair mechanisms to activate during or soon after replication, witnessed by the role of homologous, reciprocal recombination in prokaryotic repair. This view is consistent with the broad observation of recombination following DNA replication. But once this process had evolved, intercellular recombination could invade as it regulated its induction along with genes for coupling and the detection of error-prone damage. This is driven by the exploitation of the selective conditions inherently generated by background trapping. Importantly, it is not one rogue recombination gene invading at the expense of all others, but a concerted effort by the genome to synchronize times of linkage from those of dissolution. As conditions warranted, the already evolved connection between DSBs and recombination could then allow cells to recruit recombination across the genome as an assortment process via the specific creation of intergenic DSBs (see also Kuzminov 1999). Importantly, its evolution is understood from the perspective of individual genes and their association, instead of in terms of benefits to haplotypes that are destroyed by the very process itself. After coupling and recombination, the first reduction in DNA is merely to restore individuality, while the second is the ancient process of fission (Figure 3). Thus meiosis recruits heavily from the existing repair and replicative machinery.

We have long known that background trapping affects the rate of adaptation between asexual and sexual populations (Fisher 1930); this study shows a way in which it can act causally in the very creation of the populations themselves.

We warmly thank two anonymous reviewers, especially a reviewer who kindly directed our attention to supporting articles in the literature, some of which are cited herein.

LITERATURE CITED


Peck, J. R., 1994 A ruby in the rubbish: beneficial mutations, dele-
Redfield, R. J., 1988 Evolution of bacterial transformation: is sex with dead cells ever better than no sex at all? Genetics 119: 213–221.


Communicating editor: W. Stephan