“Patchy-Tachy” Leads to False Positives for Recombination

Stephanie Sun,1 Ben J. Evans,1 and G. Brian Golding *1

1Department of Biology, McMaster University, Hamilton, Ontario, Canada

*Corresponding author: E-mail: golding@mcmaster.ca.
Associate editor: Hideki Innan

Abstract

Indirect tests have detected recombination in mitochondrial DNA (mtDNA) from many animal lineages, including mammals. However, it is possible that features of the molecular evolutionary process without recombination could be incorrectly inferred by indirect tests as being due to recombination. We have identified one such example, which we call “patchy-tachy” (PT), where different partitions of sequences evolve at different rates, that leads to an excess of false positives for recombination inferred by indirect tests. To explore this phenomena, we characterized the false positive rates of six widely used indirect tests for recombination using simulations of general models for mtDNA evolution with PT but without recombination. All tests produced 30–99% false positives for recombination, although the conditions that produced the maximal level of false positives differed between the tests. To evaluate the degree to which conditions that exacerbate false positives are found in published sequence data, we turned to 20 animal mtDNA data sets in which recombination is suggested by indirect tests. Using a model where different regions of the sequences were free to evolve at different rates in different lineages, we demonstrated that PT is prevalent in many data sets in which recombination was previously inferred using indirect tests. Taken together, our results argue that PT without recombination is a viable alternative explanation for detection of widespread recombination in animal mtDNA using indirect tests.

Key words: recombination, false positives, mitochondrial DNA, heterotachy, substitution rate heterogeneity, animal mtDNA.

Introduction

There is considerable evidence to support the phenomenon of recombination between different animal mitochondrial DNA (mtDNA) molecules. Mammalian mitochondrial protein extracts can catalyze recombination (Thyagarajan et al. 1996) and mtDNA genomes may mix because mitochondria are capable of fusion (Wilson 1916; Bereiter-Hahn and Vöth 1994). Mitochondrial recombinant genotypes can be detected when two parental cells are fused (Birky 2001), and recombination products have been directly observed in cases where both paternal and maternal mtDNA were inherited; the gonads of male bivalve mussels (Ladoukakis and Zouros 2001) and in the muscle cells of a human individual (Kraytsberg et al. 2004). Although recombination has been observed in animal mtDNA, this recombination is an exception to two common assumptions in mitochondrial genetics: strict maternal inheritance and vegetative segregation. In other animals, maternal inheritance and vegetative segregation are generally thought to maintain the presence of only one mtDNA genotype in an individual (Birky 2001), a state referred to as homoplasy.

Whether mtDNA recombination is pervasive enough to require a serious reevaluation of animal population studies remains contentious due to challenges associated with collecting direct evidence of recombination which stem, at least in part, from the prevalence of homoplasy (Neiman and Taylor 2009). Due to mtDNA homoplasy, “indirect” tests for recombination have been developed wherein recombination is inferred from patterns of molecular variation as opposed to direct comparison of nonrecombined parental sequences to potentially recombined sequences. The basis of indirect tests is that over evolutionary time, occasionally more than one mtDNA genotype can be present in a cell and if recombination occurs between these molecules, it should be detectable. The molecular signature left behind by recombination could include an uneven distribution of polymorphic sites (Maynard Smith 1992; Posada and Crandall 2001), regions with high sequence similarity (Sawyer 1989), clustering of phylogenetically incompatible sites (Jakobsen and Easteal 1996), or a correlation of linkage disequilibrium with physical distance (Awadalla et al. 1999; Piganeau et al. 2004). The use of indirect recombination tests have led to reports of recombination in diverse animal mtDNA including crustaceans, amphibians (Ladoukakis and Zouros 2001), lizards (Ujvari et al. 2007), scorpions (Gantenbein et al. 2005), fish (Ciborowski et al. 2007), birds, insects, nematodes, and mammals including nonhuman primates (Maynard Smith 1992; Ladoukakis and Zouros 2001; Piganeau et al. 2004; Tsaoasis et al. 2005; White and Gemmell 2009) and humans (Kraytsberg et al. 2004). Such widespread animal mtDNA recombination raised serious concerns on the validity of clonal mtDNA inheritance, and all the models and conclusions based on this assumption.

If recombination is prevalent, clearly efforts to infer evolutionary relationships using bifurcating phylogenies are inappropriate. Evolutionary relationships are inferred using models whose parameter values are based on genetic information stored in sequences. Phylogenetic inference of a single evolutionary history (a phylogenetic tree) is confounded when different parts of the data have different evolutionary histories, which would be more accurately represented by multiple evolutionary trees. Relevant to this issue is the concept of homoplasy (not to be confused with the previously...
described homoplasy), which is when phylogenetic conflict arises due to convergent evolution. Although recurrent substitutions are clearly an important source of homoplasy in mtDNA, there lacks a broad consensus concerning the degree to which recombination contributes to phylogenetic conflict in molecular data. This raises concerns about the ability of indirect tests to distinguish between phylogenetic conflict caused by recombination and homoplasy caused by recurrent substitutions (Galtier et al. 2006).

Mutation rate heterogeneity has been suggested as a source of false positives in indirect tests for recombination (Pesole et al. 1999; Innan and Nordborg 2002; Galtier et al. 2006). Within mammalian mtDNA, mutation rates of nonsynonymous sites are highly variable from region to region (Pesole et al. 1999) and the mutation rates of specific mtDNA sites can change quickly over time, even between congeneric species (Galtier et al. 2006). Studies of mutational processes suggest that clustering of mutational events (Drake 2007) and localized elevated rates of mutational events (Wang et al. 2007; Chen et al. 2009) are frequently observed. Mutation rate heterogeneity seems to produce the most false positives when there is a high degree of contrast between rates of adjacent sites. This is supported by Innan and Nordborg (2002), where simulations with both hot and cold spots produced a significant correlation between linkage disequilibrium and distance, whereas simulations with only hot or only cold spots did not.

However, previous studies have not recovered an elevated level of false positives across diverse types of indirect recombination tests. Even with \( \theta = 200 \) and a mutation rate variation parameter \( \alpha = 0.05 \), fewer than 10% false positives were produced by Max \( \chi^2 \), GENECONV, and Reticulate (Posada and Crandall 2001). As high \( \theta \) and low \( \alpha \) are the conditions under which substitution hot spots are expected to occur, these results suggest that a simple substitution hot spot model over all sequences does not produce a substantial level of false positives. A more sophisticated substitution hot spot model over all sequences was investigated in Bruen et al. (2006) using highly correlated rates between neighboring sites and high substitution rate heterogeneity. Elevated levels of false positives were produced but only by Max \( \chi^2 \) and Reticulate and only when the site-to-site rate correlation was extremely high. Although these studies demonstrated that substitution rate heterogeneity was capable of producing false positives for recombination, their results generally suggested that most indirect tests were not susceptible to such heterogeneity under biologically realistic conditions. Therefore, the indirect evidence for widespread animal mtDNA recombination could not be discounted.

Thus, given that the mutation/substitution process can lead to unusual and varied patterns that can potentially be inappropriately inferred as due to recombination, we explored a model of substitution rate heterogeneity where subsets of the taxa might have a different rate of evolution for a portion of the data. In such a model, trees estimated from partitions share the same topology but have different branch lengths for a portion of the topology. Hereafter, we refer to this model as the patchy-tachy (PT) model. PT combines traditional models, including heterotachy (Lopez et al. 2002), which assume a single tree topology and set of branch lengths with large patches of sequences undergoing different rates of evolution across all taxa for that section. The key difference between PT and heterotachy is that in PT only a subset of the taxa (rather than all of the taxa) have accelerated or decelerated rates of evolution in a patch of the sequence.

Using simulations with varying levels of diversity, sample size, length, and other attributes, we described the effects of PT on the false positive rates of six indirect tests: GENECONV (Sawyer 1989, 1999), Max \( \chi^2 \) (Maynard Smith 1992), LD \( D^2 \) (Hill and Robertson 1968; Awadalla et al. 1999; Piganeau et al. 2004), LD \( |D^1| \) (Lewontin 1964; Awadalla et al. 1999; Piganeau et al. 2004), Reticulate (Jakobsen and Easteal 1996), and PHI (Bruen et al. 2006). Because the power of these tests has been explored in depth elsewhere (Pesola and Crandall 2001; Wiuf et al. 2001; Bruen et al. 2006), power was not investigated here. These tests are among the most powerful methods of detecting recombination when direct identification of recombinants is not feasible. They have been used to screen for recombination in a wide range of animal mtDNA.

To determine if PT might be a factor in published sequence data, we developed a simple test. This method is based on the relative likelihoods of models where the alignment is divided into partitions of various lengths. In this test, all partitions are assumed to share the same tree topology but evolutionary rates are estimated independently by partition. In this way, the relative rates of subsets of taxa in different partitions in an alignment of sequences can be compared. We used this method to test for PT in 20 animal mtDNA data sets in which recombination was detected via indirect tests. We then tested whether the mtDNA-estimated level of PT would produce an elevated rate of false positives using simulations modeled on the animal mtDNA data sets themselves. Additionally, the results of these experiments, which did not include recombination, were compared with the level of PT detected in simulations with recombination. Our results provide a reasonable, biologically feasible, alternative to the inference of widespread recombination in animal mtDNA.

Materials and Methods

Tests for Recombination

The false positive rates for six widely used (Table 1) tests of recombination were evaluated. Detailed explanations of these tests can be found in the Supplementary Material online, and in Posada and Crandall (2001). For GENECONV, we used 1,000 permutations and a gscale of 0. This test is abbreviated as GCG when both inner and outer fragments are considered and GCW when only inner fragments are used. The Max \( \chi^2 \) implementation of Piganeau et al. (2004) was used in Reticulate, sites with more than two alleles were ignored. In PHI, \( w \) was set to 100. The versions
of \(LD^2\) and \(LD^{D1}\) were the implementations used by Piganeau et al. (2004).

**Table 1. Examples of Tests Used to Detect Recombination in Animal mtDNA.**

<table>
<thead>
<tr>
<th>Test</th>
<th>Used in</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENECONV</td>
<td>Piganeau et al. (2004), Tsaousis et al. (2005), Ujvari et al. (2007),</td>
</tr>
<tr>
<td></td>
<td>Gantenbein et al. (2005), Lawson and Zhang (2009)</td>
</tr>
<tr>
<td>(LD^{D1})</td>
<td>Piganeau et al. (2004), White and Gemmell (2009), Gantenbein et al. (2005)</td>
</tr>
<tr>
<td>(LD^2)</td>
<td>Piganeau et al. (2004), Awadalla et al. (1999), White and Gemmell (2009), Gantenbein et al. (2005)</td>
</tr>
<tr>
<td>Max (\chi^2)</td>
<td>Piganeau et al. (2004), Tsaousis et al. (2005), Bruen et al. (2006), Ujvari et al. (2007), White and Gemmell (2009), Gantenbein et al. (2005), Maynard Smith and Smith (1999)</td>
</tr>
<tr>
<td>PHI</td>
<td>Bruen et al. (2006), White and Gemmell (2009)</td>
</tr>
<tr>
<td>Reticulate</td>
<td>Tsaousis et al. (2005), Bruen et al. (2006), White and Gemmell (2009), Fitzgerald et al. (1996)</td>
</tr>
</tbody>
</table>

Detecting False Recombination Signals in Nonrecombinating Simulations with PT

To investigate the effect of PT on the level of false positives, we simulated data sets without recombination but with PT such that each third of the sequence of a select clade evolves at a different rate in a closely related subset of taxa (fig. 1). The method to create these simulations is outlined in figure 2.

The creation of a simulated nucleotide data set began by generating a random tree topology in ms (Hudson 2002) with \(N\) sequences. From this tree, a clade of \(n\) sequences was chosen to be the PT clade. A copy of the tree was created and modified by scaling the internal and terminal branches of this PT clade by the factor \(pt\). The unmodified tree was used to create a partial alignment containing all sites outside the PT region. The PT tree was used to build the remainder of the alignment, which contained all sites inside the PT region. Seq-Gen (Rambaut and Grassly 1997) was used to create the nucleotide alignments. Because the unmodified and PT trees only differed in some branch lengths and not tree topology, the alignments did not include any recombination. A Jukes–Cantor model was used because at this point, we are interested in a general characterization of the effect of PT on the performance of indirect tests for recombination. Later, when we test the effect of PT on simulated animal mtDNA, we use a more appropriate model of evolution. The nucleotide data set was then screened for the total number of segregating sites, \(S\). Unless the number of segregating sites matched the target \(S\), the entire process was repeated, beginning at the generation of tree topology. The target value of \(S\) was chosen so that the simulations would reflect observed values of \(S\) from nature (specifically, a mtDNA data set from Sulawesi macaques; Evans et al. 1999).

Various PT attributes were considered including the difference in mutation rate inside versus outside the PT region, length of the PT partition in proportion to total sequence length, size of PT-affected clade, and the location of the PT partition within the alignment (specifically, in the middle surrounded by higher rate regions, or alternatively, at the periphery). These parameters and the symbols used to represent them are summarized in table 2.

**Empirically Based Simulations**

We also developed an approach to evaluate whether real data sets have PT where different rates can occur either within closely related taxa or within unrelated taxa (fig. 3). We focused on 20 animal mtDNA data sets from which recombination has been indirectly detected. These data sets included 19 of 20 data sets from the Piganeau et al. (2004) animal mtDNA data sets with the lowest probability for the null hypothesis of no recombination. *Mytilus galloprovincialis*, one of Piganeau et al. (2004)’s top 20 data sets, was excluded because recombination has been directly detected in this species (Ladoukakis and Zouros 2001). We also included mtDNA sequences of macaques collected from the Indonesian island of Sulawesi (Evans et al. 1999). These sequences are from seven macaque species: *Macaca nigra*, *M. nigrescens*, *M. hecki*, *M. tonkeana*, *M. ochreata*, *M. brunnescens*, and *M. maura*.  

![Image](https://example.com/image.png)  

**Fig. 1.** Example of five partitioned models for \(N\) sequences of \(L\) nucleotides. Partition A is represented by solid lines, partition B is represented by broken lines. Partitions A and B share the same topology but branch lengths of each section are estimated independently.
The lnL max for partition A and the lnL max for partition B were summed to obtain the total likelihoods with alternative partition regimes. Using the Akaike information criterion (AIC; Akaiake 1974), the null model (AAA) and PT models (ABA, BAA, AAB, or B−A+B−; fig. 1) were compared. The AIC was calculated as $2(df) - 2\ln L$ where there are 9 df in the AAA model (5 rate ratios, 3 nucleotide frequencies, and 1 $\Gamma$ shape parameter), and 18 in all other models (9 df for partition A and 9 df for partition B). AIC weights were calculated as in Wagenmakers and Farrell (2004) and represent the relative likelihood of the model being the best model.

The GTR + $\Gamma$ + codon parameters and partition pattern of the “best partitioned model” were used to create 1,000 partitioned simulations in Seq-Gen (i.e., one Seq-Gen run per partition). The length, sample size, and S of the simulations were matched to the real data sets analyzed by Piganeau et al. (2004) (supplementary table 1, Supplementary Material online). Partitions A and B used the same MrBayes consensus topology thereby ensuring recombination was not included. Sequences from partitions A and B were then concatenated according to the favored partition regime and tested for recombination. Null (no-PT, no recombination) AAA simulations were created in a similar manner except that only a single set of ML parameters and a single Seq-Gen simulation was used.

**Testing Animal mtDNA for PT**

There are two criteria in the test for PT: 1) The data favor a PT model (see fig. 1) rather than the null AAA model and 2) branch lengths estimated from partition A must be significantly different from the corresponding branch lengths in partition B. The type of partitioned model favored by the data was determined using the AIC as described above.

First, branch lengths were estimated from each partition. A distribution of null branch length ratios was calculated by partitioning the null AAA simulations into sections according to each PT model and then dividing branch lengths of corresponding branches from each section. A distribution of observed branch length ratios under the PT model was calculated by dividing each branch length from partition A by the corresponding branch length from partition B. Branches where the rate in partition A equals the rate in partition B would have a branch length ratio of 1. Branches with a large rate difference between partition A and partition B would deviate further from 1. Each branch length ratio from the PT model was then compared with the null distribution of branch length ratios from the null (AAA) model and a P value was calculated. These P values represented the proportion of null branch length ratios that were equally or more extreme than the simulated PT branch length ratio. P values less than or equal to 0.05 were marked as possible signals for PT. False discovery rate (FDR; Benjamini and Hochberg 1995) was used to correct the number of potential PT signals for multiple tests.
Testing Simulated Data with Recombination for PT

It is also possible that true recombination events might lead to PT. Therefore, further simulations were carried out to quantify such signals that might be generated in association with recombination. Two levels of recombination were tested. In the first, biologically feasible recombination rates based on autosomal DNA and mutation rates based on mtDNA ($c$ and $\mu$, respectively) were used. We used a value of $\mu$ derived from great apes of $5.2266 \times 10^{-7}$ mutations/site/generation (Lynch 2007). To calculate the recombination rate, $c$, we assumed the same ratio of recombination to mutation per nucleotide site as in humans, $\frac{c}{\mu} = 0.6$ (Ptak et al. 2004; Lynch 2007). Using these values, we calculated $c$ to be $3.14 \times 10^{-7}$ events/site/generation. Other estimates of primate autosomal recombination rates have been as low as $c = 1.2 \times 10^{-8}$ with $\mu = 2 \times 10^{-8}$ (Becquet and Przeworski 2007). This suggests that our estimate for autosomal recombination rate is high. To test an extreme bound for $c$ ($3.14 \times 10^{-6}$), we increased the autosomal recombination rate by a factor of 10. This second $c$ was intentionally set extremely high to emphasize any possible effect of recombination on PT detection ($\mu$ was left unchanged). Each model was tested with 100 simulations that were generated using RECODON (Arenas and Posada 2007). Simulations were tested for PT in the same manner as described in the section “Testing Animal mtDNA for PT.” Reticulate, PHI, and Max $\chi^2$ were used to measure the level of detectable recombination.

Results

Effect of PT on Recombination False Positives

The proportion of simulated data sets with no recombination that falsely detect recombination when the simulations have PT is shown in figure 4. When $N = 15$, the level of false positives detected by Reticulate ranged from 0.07% to 67%, by PHI ranged 0.04–34%, by GENECONV ranged 4–99.7%, by Max $\chi^2$ ranged 3–99.9%, by LD$^2$ ranged 4–75%, and by LD$|D'|$ ranged 4–72%. The level detected by Reticulate ranged from 0.07% to 72%, by PHI ranged 0.01–10%, by GENECONV ranged 4–99.8%, by Max $\chi^2$ ranged 4–99.9%, by LD$^2$ ranged 4–52%, and by LD$|D'|$ ranged 1–99.6%. Generally, in simulations...
with PT, the likelihood of a test reporting a false positive increased as the scaling factor for clade $n$ became more extreme. This effect was particularly pronounced when the sequences contained a high proportion of polymorphic sites. When the region with PT sites is located at the edge of the alignment, the level of false positives of PHI, $L D_r^2$, and $L D |D'|$ decreases to below 10% but Reticulate, GENECONV, and Max $\chi^2$ false positives for recombination were unaffected (supplementary fig. 1 vs. supplementary fig. 3; supplementary fig. 2 vs. supplementary fig. 2, Supplementary Material online). Shortening the overall sequence length results in fewer false positives for recombination (supplementary fig. 1 vs. supplementary fig. 5; supplementary fig. 2 vs. supplementary fig. 6, Supplementary Material online). Generally, the greatest fluctuation in the number of false positives between repeated simulations was $\sim$2%. In simulations without PT (a scaling factor of 1), the tests correctly returned $\sim$5% false positives.

The rate of evolution influenced whether scaling branch lengths of a subset of the taxa or of all taxa in the simulations gave results with more false positives for recombination. With a moderate mutation rate, the level of false positives for recombination is lower when the substitution rate of a partition is scaled in all sequences rather than only a subset of sequences (table 3a and 3b). However, the opposite is true when the mutation rate is high (table 3c).

### Detecting PT in Animal mtDNA

Three of the 20 animal data sets tested did not have detectable PT. *Mandrillus sphinx* favored a nonpartitioned AAA model over a partitioned one (table 4), whereas *Dendroica petechia* (Warbler) and *Macrodon ancylodon* (King weakfish) were found not to possess PT after the FDR correction was applied. If we consider 5% or less as an acceptable level of false positives, *Macrodon ancylodon* and, according to GENECONV and $L D_r^2$, *Mandrillus sphinx* and *Dendroica petechia* did not have detectable PT but did have elevated false positives for recombination based on simulations without recombination using evolutionary parameters derived from this data set (table 5). All other data sets had PT and also had an elevated false positive recombination rate, although the degree of false positives ranged from 99% to 7% (*Microtus longicaudus* [Vole] to *Merlangius merlangus* [Whiting]).

### Table 3. Effect of Cold Clade Size ($n$) with Different Background Mutation Rates.

<table>
<thead>
<tr>
<th>$n$</th>
<th>RET</th>
<th>PHI</th>
<th>GCG</th>
<th>GCI</th>
<th>MX</th>
<th>LDR</th>
<th>LDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>8</td>
<td>6</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>7</td>
<td>38</td>
<td>35</td>
<td>35</td>
<td>68</td>
<td>13</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>6</td>
<td>17</td>
<td>9</td>
<td>17</td>
<td>38</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n$</th>
<th>RET</th>
<th>PHI</th>
<th>GCG</th>
<th>GCI</th>
<th>MX</th>
<th>LDR</th>
<th>LDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>11</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>7</td>
<td>89</td>
<td>88</td>
<td>69</td>
<td>83</td>
<td>17</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>5</td>
<td>31</td>
<td>22</td>
<td>10</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n$</th>
<th>RET</th>
<th>PHI</th>
<th>GCG</th>
<th>GCI</th>
<th>MX</th>
<th>LDR</th>
<th>LDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>66.6</td>
<td>34.0</td>
<td>99.7</td>
<td>99.7</td>
<td>99.9</td>
<td>75.3</td>
<td>99.7</td>
</tr>
<tr>
<td>10</td>
<td>23.3</td>
<td>6.9</td>
<td>99.8</td>
<td>99.8</td>
<td>98.9</td>
<td>95.4</td>
<td>54.2</td>
</tr>
<tr>
<td>15</td>
<td>63.6</td>
<td>17.5</td>
<td>99.7</td>
<td>99.7</td>
<td>99.9</td>
<td>48.0</td>
<td>99.9</td>
</tr>
</tbody>
</table>

**Note.** — $N = 15, L = 1,200$ bp, $p =$ middle, $l = 800$ bp.


### Table 4. AIC Values for Partition Models.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>AIC Weights</th>
<th>Data Set</th>
<th>AIC Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sulawesi macaques</td>
<td>AAA AAB</td>
<td>Omphiocephalus hodgsoni</td>
<td>AAA AAB</td>
</tr>
<tr>
<td></td>
<td>7.590 × 10⁻⁶⁵ 0.999</td>
<td>(springtail) 2.262 × 10⁻³ 0.722</td>
<td></td>
</tr>
<tr>
<td>Bursaphelenchus conicaudatus (nematode)</td>
<td>AAA AAB</td>
<td>Alpheus lotti</td>
<td>AAA B A+B</td>
</tr>
<tr>
<td>Microtus salmoides (bass)</td>
<td>AAA AAB</td>
<td>Macrodon anglydon</td>
<td>AAA AAB</td>
</tr>
<tr>
<td>Macaca nemestrina (pig-tailed macaque)</td>
<td>AAA AAB</td>
<td>Papio papio</td>
<td>AAA BAA</td>
</tr>
<tr>
<td>Microtus longicaudus (vole)</td>
<td>AAA AAB</td>
<td>Camphylorhynchus bruneicapr</td>
<td>AAA ABA</td>
</tr>
<tr>
<td>Vesicomya pacifica (bivalve)</td>
<td>AAA B A+B</td>
<td>Bradypodion occidentale</td>
<td>AAA AAB</td>
</tr>
<tr>
<td>Mandrillus sphinx (mandrill)</td>
<td>AAA ABB</td>
<td>Passerella iliaca</td>
<td>AAA AAB</td>
</tr>
<tr>
<td>Libellula quadrimaculata (dragonfly)</td>
<td>AAA ABA</td>
<td>Merlangius merlangus</td>
<td>AAA ABA</td>
</tr>
<tr>
<td>Dendroica petechia (warbler)</td>
<td>AAA ABA</td>
<td>Gonatus onyx</td>
<td>AAA ABA</td>
</tr>
<tr>
<td>Apodemus sylvaticus (woodmouse)</td>
<td>AAA AAB</td>
<td>Grus antigone</td>
<td>AAA B A+B</td>
</tr>
</tbody>
</table>

NOTE.—AIC weights for the no partition model (AAA) and the best partitioned model. AIC weights represent the relative likelihood of the model being the best model. Data sets are ordered by strength of evidence for recombination according to Piganeau et al. (2004).

according to GENECONV; table 5). It is interesting to note that data sets with stronger evidence for recombination as reported in Piganeau et al. (2004) also had parameters estimates that produced nonrecombining, PT simulations with the highest levels of false positives. In summary, 17 out of the 20 data sets possess both detectable PT and had parameter estimates that produced an elevated level of false positives for recombination in simulations without recombination.

### Effect of Recombination on the Incidence of PT

The mean number of recombination events using the autosomal rate of recombination, c, was 6 events/tree. In simulations generated using the extreme c, the mean number of recombination events was 58 events/tree (fig. 5). Although all simulations were recombinant, only 8% of them had detectable recombination under the autosomal recombination model based on the indirect tests we examined. Of these 8%, PT was detected in 25% (table 6). Under the extreme recombination model, 27% of simulations had detectable recombination and of these, PT was detected in about 52%.

### Discussion

Effect of PT on the Level of Recombination False Positives Using simulations under a simple Jukes–Cantor and a
model of PT within a subset of closely related clades, we demonstrated the severe degree to which PT can magnify the level of false positives detected by indirect tests for recombination. These simulations demonstrate the relative performance of indirect tests given a particular set of simulations with PT in a subset of related taxa. The substitution distribution tests (GENECONV, Max $\chi^2$) performed the worst, followed by LD$|D'|$, Reticulate, LDr$^2$, and PHI, respectively. Two factors influence the magnitude of this PT effect: the number of segregating sites and the degree of scaling between partitions of sites.

Performance of Substitution Distribution Tests for Recombination

These tests assume that there is an even distribution of segregating sites when recombination is not present. This assumption is violated by PT, which causes substitutions to be unevenly distributed across and between sequences. The GENECONV criteria of considering only inner fragments (GCI) is considered more conservative than the alternative of considering all global fragments detected (GC), however, GCI finds the same high level of false positives as GC (supplementary figs. 5–7, Supplementary Material online).

Other substitution distribution methods include the Homoplasy Test (Maynard Smith and Smith 1998), Informative Sites Test (PIST; Worobey 2001), Chimaera (Posada and Crandall 2001), the Runs Test (Takahata 1994), and the Sneath Test (Sneath 1995). These substitution distribution tests will probably also produce elevated false positives in a PT model but have not been tested here. This is especially true considering the performance of Max $\chi^2$ and GENECONV, as they are the most powerful and robust of the substitution distribution methods. Indeed, two tests, the Homoplasy Test and PIST, have been found to produce high levels of false positives with extreme levels of rate variation (Posada and Crandall 2001).

Performance of Linkage Disequilibrium and Distance Tests for Recombination

Although LDr$^2$ is a more powerful test than LD$|D'|$ to detect recombination when it is present (White and Gemmell 2009), LDr$^2$ found many more false positives than LD$|D'|$, suggesting that the gain comes with a cost in accuracy. The null hypothesis that without recombination, linkage disequilibrium does not correlate with distance fails because clusters of unusually fast or unusually slow-evolving sites (relative to the rest of the alignment) can produce a negative correlation between linkage disequilibrium and distance (Innan and Nordborg 2002). The major difference between LD$|D'|$ and LDr$^2$ is that LDr$^2$ can only measure linkage disequilibrium when all four genotypes are present (the two parentals and two recombinants), and when allele frequencies are moderate to high (Awadalla et al. 2000). It is not surprising then, that LDr$^2$ has a higher level of false positives than LD$|D'|$, although LD$|D'|$ will still produce an elevated level of false positives under certain conditions.

Performance of Phylogenetic Compatibility Tests for Recombination

Reticulate and PHI are compatibility methods and of all the tests evaluated here, they are the most robust to mutation cold spots (but still produced up to 30% false positives). The improved false positive rate of the most robust test, PHI, comes at a price. PHI tends to be overly conservative when there are too few informative sites or too few

Table 6. Recombination and PT in Recombined Simulations Compared with the PT Simulation.

<table>
<thead>
<tr>
<th>Recombination detected</th>
<th>PT Detected</th>
<th>Simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High $c$</td>
<td>Extreme $c$</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>61</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>31</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>6</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>2</td>
</tr>
</tbody>
</table>

Note.—Percentage of simulations in which recombination was detected with Reticulate, PHI, or Max $\chi^2$ and the number of simulations in which PT was detected. High: autosomal recombination rate ($c = 3.14 \times 10^{-7}$; 100 simulations); Extreme: $c = 3.14 \times 10^{-6}$; 100 simulations; mtDNA: no recombination animal mtDNA simulations $c = 0$, 1,000 simulations, recombination is considered detected in data sets with over 5% false positives.
incompatibilities; this tends to occur when alignments have fewer than 15 sequences or when nucleotide diversity is below 5% (Bruen et al. 2006).

PT Versus Mutation Rate Heterogeneity
Although PT and rate heterogeneity share similarities, they are not the same phenomena. Rate heterogeneity, as commonly applied, refers to variation in substitution rates between sites across all sequences (i.e., \( n = N \)), whereas PT refers to variation in substitution rates between subsets of sites and in a subset of sequences (i.e., \( n \neq N \)). PT is capable of producing an elevated level of false positives which cannot be accounted for by rate heterogeneity alone (table 3a and 3b). Likewise, rate heterogeneity produces an elevated level of false positives which cannot be accounted for by PT alone (table 3c).

In summary, LD/\( D^{1} \), Reticulate, and PHI are comparatively less susceptible to false positives by PT than GENECONV, Max \( \chi^{2} \), and LD/r\( 2 \). When LD/\( D^{1} \), Reticulate, and PHI do falsely infer recombination, it is probably because sites have mutated multiple times in such a way that all four genotypes are present at a site (LD/\( D^{1} \)) and/or informative sites outside the cold region appear phylogenetically incompatible with sites inside the cold region (Reticulate and PHI). This would explain why simulations with a small cold clade but a large cold region produced the most false positives for LD/\( D^{1} \), Reticulate, and PHI; a large cold region forces the few remaining noncold sites to accept more mutations per site to maintain the same overall rate of evolution, but if too many clades contain the cold region, there will be too few informative sites to detect any signal.

Doubling the number of sequences in the data set from 15 to 30 did not substantially improve the performance of Max \( \chi^{2} \), GENECONV, LD/\( D^{1} \), or Reticulate but improved the performance of LD/r\( 2 \) and PHI (fig. 4). This contradicts a previous study, where an increase in number of sequences from 10 to 50 led to an increase in level of false positives detected from 10% to over 50% in PHI, Reticulate, and Max \( \chi^{2} \) (Bruen et al. 2006). This could be due to additional factors included in the previous model, including exponential growth, extreme site-specific rate heterogeneity, and/or the method of simulating mutation hot spots, which would be shared across all sequences.

PT in Animal mtDNA
Using simulations that can detect PT in either closely related or nonclosely related lineages, we explored the possibility that PT exists in animal mtDNA that was previously reported to carry a signature of recombination. We asked, how prevalent PT is in mtDNA that have had recombination reported and if that degree of PT is sufficient to generate false positives for recombination. PT was detected in 85% of the animal mtDNA data sets and of these, all had an elevated level of false positives for recombination (table 5). The three data sets that did not have detectable PT were Mandrillus sphinx, Dendroica petechia, and Macrodon ancyodon.

Ninety-five percent of the animal mtDNA data sets with previously reported recombination favored a partitioned model over the nonpartitioned AAA model (table 4). Without exception, tests produced more recombination false positives when the data were partitioned than when only one set of parameters and branch lengths was provided for the whole sequence. Admittedly, our method of partitioning is a crude manner to detect PT when modeling empirical mtDNA data sets. Nonetheless, this relatively coarse PT model fits many of these data sets significantly better than a model without PT (see below). An interesting direction for future work would be the development of a more general PT model where the sample size and length of the PT region are estimated from the data.

In Piganeau et al. (2004)’s survey of recombination in 267 animal mtDNA, only four data sets were statistically significant—Bursaphelenchus conicaudatus, Macaca nemestrina, Microtus longicaudus, and Micropterus salmoides. We found that, under a PT model, these four “most probable” recombining data sets were consistently among those producing the highest level of false positives across numerous tests (table 5).

Effect of Recombination on the Incidence of PT
These simulations offer an opportunity to compare the level of PT detected in animal mtDNA to the level expected due to recombination. Interestingly, recombination is largely undetected in simulations with recombination that were simulated using a high autosomal estimate of \( c \) and even an extreme artificially high \( c \) (table 6). This may be because only Reticulate, PHI, and Max \( \chi^{2} \) were used to test the recombined simulations. Nevertheless, when we compare the recombined simulations to the animal mtDNA using Reticulate, PHI, and Max \( \chi^{2} \); only, we can compare the level of PT observed in the animal mtDNA to the expected level due to recombination alone.

Of the 8% of autosomal level recombination simulations with detectable recombination, 25% of these possess PT (table 6). Of the 27% of extreme \( c \) simulations with detectable recombination, about half (52%) have PT. In contrast, in the animal mtDNA data sets, 80% of simulations have detectable recombination (using Reticulate, PHI, or Max \( \chi^{2} \) only), and of these, 94% have PT. This suggests that either the PT detected in the animal mtDNA is due to a phenomenally high recombination rate, much higher than the extreme rate tested here, or, that the excess 42–69% of PT in animal mtDNA cannot be explained by any biologically feasible recombination rate.

Conclusion
The possibility of widespread recombination in animal mtDNA caught immediate attention as it implied important consequences for phylogenetic and population studies. We show here that a specific type of mutation heterotachy, which we call PT, is capable of producing extremely high levels of false positives for recombination in indirect tests. We present a method of measuring PT, and using this method
on previously reported recombining animal mtDNA, show that the level of PT present will produce elevated levels of false positives in almost all mtDNA data sets tested. Finally, we also demonstrate that the level of PT measured in the animal mtDNA cannot be explained by recombination alone. These results do not refute the possibility that recombination can and does occur in some animal mtDNA. Rather, it casts doubt on the ability of indirect recombination tests to distinguish between recombination and PT. This finding is consistent with previous studies which suggest that heterotachy is widespread in the mitochondrial genome; at least 28–95% polymorphic sites are heterotachous in mitochondrial coding regions, and the position of heterotachous sites does not appear to be tied to functional divergence nor to spatial structure of the protein (Lopez et al. 2002).

The indirect tests evaluated in this study have been used in studies screening a much wider range of animal mtDNA than we studied here (Tsaoasis et al. 2005; Ujvari et al. 2007). It seems probable that PT in these other data sets could also produce elevated false positives. Other indirect tests that have been used to detect recombination do not consider PT and could produce as many or more false positives as have been reported here.

**Supplementary Material**

Supplementary figures, and tables are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

**Acknowledgments**

The authors thank Dr Gwenaëil Piganeau who generously shared her LD|\(d^2\), LD|\(D^1\), and Max \(\chi^2\) scripts and provided assistance in using them. Anonymous reviewers put a great deal of effort in improving this manuscript and we sincerely thank them for their efforts. This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (www.sharcnet.ca) and was supported by Natural Sciences and Engineering Research Council discovery grants to G.B.G. and B.J.E.

**References**


