SEQLinco: A Novel Approach to Use Sequence Data for Linkage Analysis

Gao T. Wang1, Di Zhang1, Biao Li1, Hang Dai1, Suzanne M. Leal1, \*

1Center for Statistical Genetics, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA

**Methods**

***Combined Haplotype Pattern coding***

We propose a *Combined Haplotype Pattern* (CHP)method to …

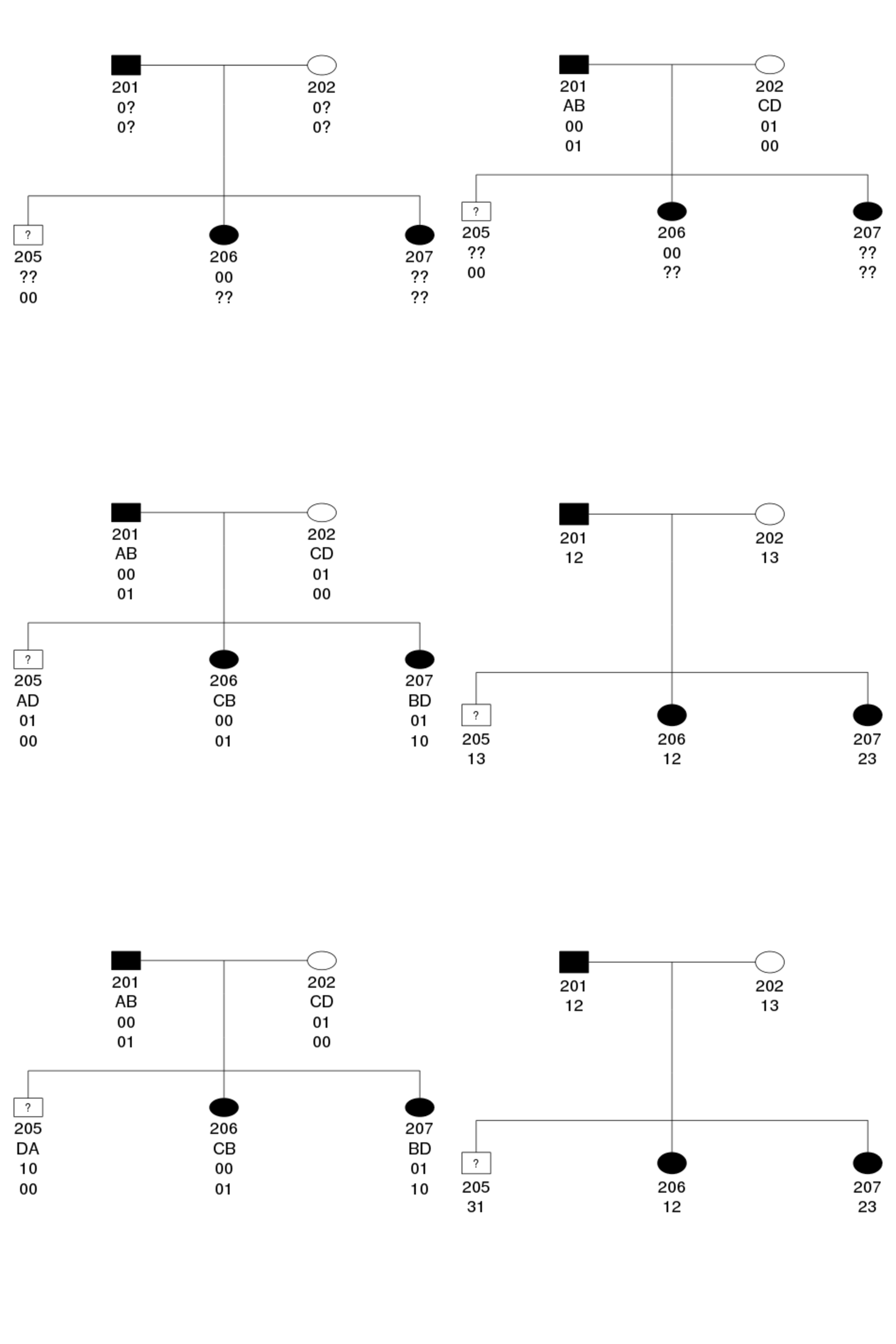
***Haplotype reconstruction***

Rules:

1. For heterozygous sites, always place 1 to the right side if it is the first 1 being placed to the haplotype. For upcoming 1s in offspring, check for Mendelian conflict before placing them in and if there are still multiple choices, place 1 to the right side. For upcoming 1s in parents, place them to the right and check for Mendelian conflict in all offspring; if found, place it on left and check again.
2. If there are any Mendelian conflicts that cannot be resolved by placing 1 on either side, throw away the site.

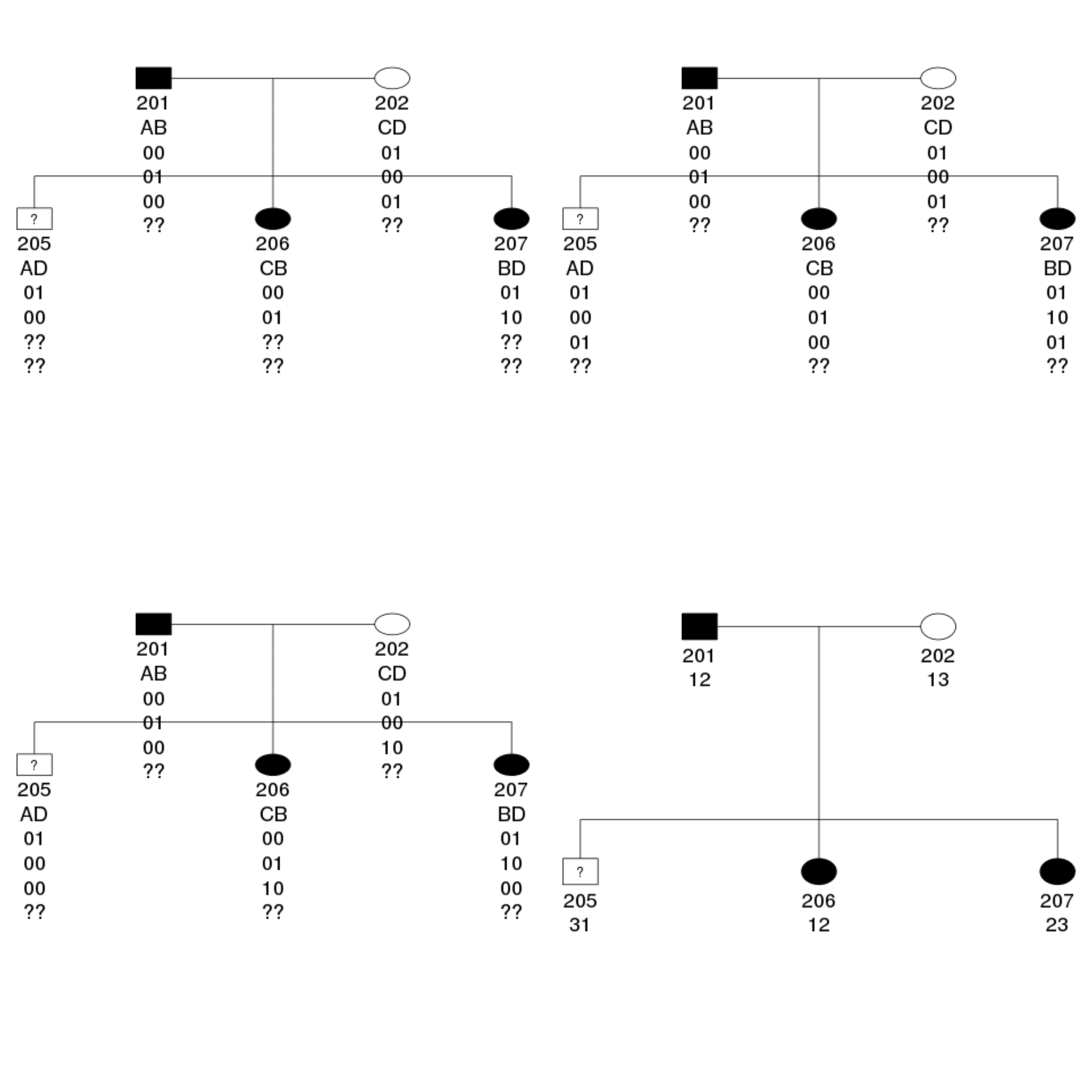
First we work on sites where any offspring is 00 or 11:

1. *Resolve parental sites using 00&11 sites in offspring*: For sites where any offspring is 00, configuration of that site on both parents should be 00, 01 or 10. Given knowledge of parental genotype we apply Rule 1 to determine the parental haplotypes for that site. The same applies for 11 sites in offspring.
   1. Figure 1.1: From site 2 in #205 and site 1 in #206 we can determine the left side of all sites in #201 and #202.
   2. Figure 1.2: Determination of the right side of all sites in #201 and #202 are straightforward given knowledge of parental genotype.
2. *Determine more sites on offspring using parents*: For sites where parental haplotypes are all available due to previous step, we use Rule 1 to determine offspring haplotypes on those sites based on parental haplotypes and offspring genotypes.
   1. Figure 1.3 and 1.4: For site 1 in #205 and site 2 in #206 by Rule 1 we determine the haplotypes as 01. For site 1 in #207 by Rule 1 we place 01, and also by Rule 1 we check for Mendelian conflict and determine the haplotype as 10. We then apply CHP*.*
   2. Figure 1.5 and 1.6: What if we do not apply Rule 1 and use a different arbitrarily picked configuration for an offspring? For example for site 1 in #205 we use another configuration. The CHP result is not effected.



We then move on to sites where no offspring is 00 or 11:

1. *Resolve sites with parental 00 or 11*: For sites where one parent is 00, we place the site on the other parent by Rule 1, and determine the corresponding sites for offspring by Rule 1.
   1. Figure 2.1: We consider two more sites (site 3 and site 4) in addition to the sites already determined in Figure 1. From site 3 in #201 we write site 3 in #202 as 01.
   2. Figure 2.2 and 2.3: We determine offspring haplotypes on site 3. It is either wrong on #206 or wrong on #205&#207. Rule 1 cannot resolve it, so we discard it by Rule 2.
2. *Discard sites with everyone being heterozygous*: Such sites are uninformative even if resolved (FIXME: examples).
   1. Figure 2.4: We discard site 4. Basically we ignored site 3 and 4 and end up the same CHP as Figure 1.



***Recombination events***

Consequence: a series of Mendelian conflicts after certain site.

***Missing data***

**Figure legends**

**Figure 1.**

**Appendix**