Predicting Chromosomal Locations of Genetically Mapped Loci in Maize Using the Morgan2McClintock Translator

Carolyn J. Lawrence,*† Trent E. Seigfried,* Hank W. Bass,‡ and Lorinda K. Anderson§

*USDA-ARS, Corn Insect and Crop Genetics Research Unit, Iowa State University, Ames, IA 50011; †Department of Agronomy and Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA 50011; ‡Department of Biological Science, Florida State University, Tallahassee, FL 32306-4370; and §Department of Biology, Colorado State University, Fort Collins, CO 80523
RUNNING HEAD
Genetic to cytological map converter

KEYWORDS
Genetics, cytogenetics, maps, maize, bioinformatics.

CORRESPONDING AUTHOR
Carolyn J. Lawrence
1565 Agronomy Hall
Iowa State University, Ames, IA 50014
Telephone: (515) 294-7380
FAX: (515) 294-2299
E-mail: triffid@iastate.edu.
ABSTRACT

The Morgan2McClintock Translator permits prediction of meiotic pachytene chromosome map positions from recombination-based linkage data using recombination nodule frequency distributions. Its outputs permit estimation of DNA content between mapped loci and help to create an integrated overview of the maize nuclear genome structure.
Two fundamentally different but collinear types of gene maps can be produced, linkage maps and physical maps. Classical linkage (genetic) maps are based on allele-recombination frequencies, whereas physical maps are based on the linear DNA molecules that comprise the chromosomes.

In maize, a model genetic and major agricultural species, over 1200 high-resolution linkage maps composed of thousands of markers are available, whereas detailed physical maps of DNA sequence and chromosome structure are still in development. The three main types of maize physical maps differ in level of molecular resolution. They are (1) genome sequence assembly maps, at DNA base-pair resolution (see, e.g., Dong et al. 2005; Fu et al. 2005); (2) fingerprint-contig (FPC) maps, resolved at the level of overlapping restriction fragments from cloned segments of genomic DNA (see, e.g., Pampanwar et al. 2005); and (3) cytological maps constructed by microscopic observation of pachytene chromosome structure (e.g., the Cytogenetic FISH 9 map created by Koumaris and Bass 2003 and Amarillo and Bass 2004).

Linkage and physical maps have different coordinate systems for positioning loci. The genetic map unit is called a "centiMorgan" (cM) in honor of Thomas Hunt Morgan. One cM is equal to 1% crossing over between two linked loci. Fingerprint-contig and genomic-assembly maps are measured in base pairs, whereas physical maps based on pachytene chromosome structure (also called cytological or cytogenetic maps) position each locus as the fractional distance along the arm from the centromere to the telomere. Recently, maize researchers have begun to call the unit of this sort of map denomination a "centiMcClintock" (cMC) in honor of maize genetics pioneer Barbara McClintock. Here we formally define 1 cMC as 1% of the length of the chromosome arm upon which
a given locus resides. For example, if the short arm of chromosome 9 is 8.70 µm in length and the bronze1 (bz1) locus lies 5.66 µm from the centromere on that chromosome arm, bz1 lies \( \frac{5.66}{8.70} \times 100 = 65\% \) of the distance from the centromere to the chromosome tip or 65 cMC from the centromere. A locus at position 66 would lie exactly 1 cMC from the bz1 locus. Because maize chromosome arm lengths vary and the centiMcClintock is a relative unit, 1 cMC on, e.g., the short arm of chromosome 9 does not necessarily consist of the same number of micrometers as 1 cMC on any of the 19 other chromosome arms. The cytological conventions are further described and defined at http://www.maizegdb.org/coordinateDef.php.

Recombination rates vary tremendously along individual chromosomes such that the map distance between two loci on a linkage map may not accurately predict the physical distance between them (Anderson et al. 2004). This variation has made integrating the two types of maps difficult and also has important implications for genome-assembly efforts and positional-cloning strategies (Sadder and Weber, 2002).

A method for linking genetic maps with chromosome structure has recently been developed. Anderson et al. (2003) determined the frequency distributions of recombination nodules (RN) along the 10 pachytene chromosomes of maize. Because each RN represents a crossover on the physical structure of the chromosome, these RN maps are unique in that they contain both linkage and cytological information that allows the prediction of the cytological position of any genetically mapped marker (Anderson et al. 2004). We have developed a tool, the Morgan2McClintock Translator (accessible at http://www.lawrencelab.org/Morgan2McClintock), which automates the cytological-position prediction process for any input linkage data.
Conversion of maize linkage map coordinates into cytological coordinates requires both linkage data and RN frequencies as input. The Morgan2McClintock Translator includes as data files the maize RN map (Anderson et al. 2003) as well as two genetic maps, the UMC 1998 map (Davis et al. 1999) and the 1997 Genetic map (Neuffer et al. 1995). More than a thousand other genetic maps, which also can be used as infiles, are available at MaizeGDB (Lawrence et al. 2005 and http://www.maizegdb.org/map.php). The translator itself was coded with PHP, and the equations it uses to convert linkage maps into cytological maps are those described by Anderson et al. (2004). The application can be run online, or it can be downloaded for local use on any machine equipped to serve PHP. Aspects of the input and output displays for the translator are shown in Figure 1, for the UMC 98 genetic map (Davis et al. 1999).

The distribution of RNs provides an important connection between genetic maps and chromosomal structure, which has allowed the examination of gene distribution at the chromosomal level in maize (Anderson et al. 2005). This integration also permits estimation of DNA and chromosomal distances between genetic loci, a feature that will assist in the sequence assembly of the maize genome. Theoretically, this approach is applicable to other organisms with comparable cytological crossover-distribution data such as tomato (Sherman and Stack 1995) and mouse (Froenicke et al. 2002), and we plan to develop a set of similar tools for these organisms that should be useful in comparing genetic and chromosomal aspects of genomes in different species.

Use of the maize Morgan2McClintock Translator will allow researchers to integrate previously disparate views of maize genome structure. For example, the maize
cytological maps (http://www.maizegdb.org/cgi-bin/displaycompletemaprecord.cgi?id=40028) are predominantly annotated with chromosomal translocation breakpoints (Coe 1994). For most breakpoints, corresponding germplasm is available from the Maize Genetics Cooperation—Stock Center (Scholl et al. 2003). Integrating the cytological breakpoint positions with genetic linkage maps would enhance the application of available translocation stocks to genome research, breeding programs, and chromosome engineering efforts. This is one among many ways in which the Morgan2McClintock Translator could be used specifically to add value to maize genetics and structural genomics research and more generally to aid in meiotic chromosome research.
We thank Nigel Walker and Anne B. Thistle for critical reading of the manuscript. This work was supported by the USDA-ARS and by the National Science Foundation (DBI-0321639 to H.W.B. and MCB-314644 to L.K.A.).
LITERATURE CITED


Sherman, J. D. and S. M. Stack, 1995 Two-dimensional spreads of synaptonemal
complexes from solanaceous plants. VI. High-resolution recombination nodule map for tomato (*Lycopersicon esculentum*). Genetics **141**: 683–708.
FIGURE LEGEND

FIGURE 1.—The Morgan2McClintock Translator. Screen capture images taken from http://www.lawrencelab.org/Morgan2McClintock show examples of data input (upper panel) and output (lower panel). (A) The user first chooses the maize linkage group as chromosome number (arrow at Step 1) and then the corresponding centiMorgan linkage-map data set (arrow at Step 2). The linkage map data can be chosen from among stored data sets available for common maps, or it can be pasted directly into a text box for map data not currently stored. Clicking the “Calculate” button submits input data and calculates centiMcClintock values from the RN frequency distribution. The output web page contains a table that summarizes one locus per row and includes columns that describe the input data in centiMorgans (B) and the output data in predicted locations along the pachytene chromosome, expressed in microns and in centiMcClintocks (C).
Morgan2McCIntock Translator

This tool uses the maize Recombination Nodule map (Anderson et al. 2003 and 2004) to calculate approximate chromosomal positions for loci given a genetic map. To run the calculator on your own machine, visit the download page.

**Step 1:** Select a chromosome: 1

**Step 2:** Select a map to use...

<table>
<thead>
<tr>
<th>UMC 98 (genetic)</th>
</tr>
</thead>
</table>

... or paste your genetic map in

How To Submit Your Own Map Data:
Your map data MUST include the following:
(1) Your map must consist of a number of lines. Each line should start with the name of the locus, followed by a tab, followed by the centiMorgan value for that locus.
(2) The program assumes that the tip of the short arm is at 0.0 cM UNLESS you submit a locus with a cM value of less than zero.
(3) The program also assumes that the tip of the long arm is at the cM value of your highest submitted cM value. So, if you believe that the tip is not exactly at this locus, you should add a tip1 locus at the end with the cM value of where you believe the tip to be.
(4) There must be ONE locus beginning with the characters "cent". The program will assume that this is the centromere.

Morgan2McCIntock Translator

The table below shows your input in the first two columns. The third column shows the relative position of each genetic locus as a fraction of map length from the centromere. This data is used to adjust the genetic map length to fit the RN-cM map (fourth column). The last three columns show the predicted cytological location of the locus on the chromosome.

To access the maize RN map raw data and related information, visit MaizeGDB.

**Map:** UMC 98
**Chromosome:** 1

<table>
<thead>
<tr>
<th>Locus</th>
<th>centiMorgan (cM)</th>
<th>As fraction of cM map from centromere</th>
<th>Converted to RN-cM</th>
<th>Corresponding absolute position on SC/chromosome (μm from tip of short arm)</th>
<th>Position as fractional length of arm from centromere (centiMcCIntocks)</th>
<th>Arm</th>
</tr>
</thead>
<tbody>
<tr>
<td>csu804b(dnp)</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.0</td>
<td>1.00 (100)</td>
<td>S</td>
</tr>
<tr>
<td>rgpc654</td>
<td>6.40</td>
<td>0.95</td>
<td>3.16</td>
<td>0.8</td>
<td>0.96 (96)</td>
<td>S</td>
</tr>
<tr>
<td>tub1 (CBM 1.01)</td>
<td>11.00</td>
<td>0.91</td>
<td>5.42</td>
<td>1.0</td>
<td>0.95 (95)</td>
<td>S</td>
</tr>
<tr>
<td>csu738</td>
<td>11.00</td>
<td>0.91</td>
<td>5.42</td>
<td>1.0</td>
<td>0.95 (95)</td>
<td>S</td>
</tr>
<tr>
<td>umc94a</td>
<td>11.90</td>
<td>0.90</td>
<td>5.87</td>
<td>1.0</td>
<td>0.95 (95)</td>
<td>S</td>
</tr>
<tr>
<td>bnl8.05a</td>
<td>12.00</td>
<td>0.90</td>
<td>5.92</td>
<td>1.0</td>
<td>0.95 (95)</td>
<td>S</td>
</tr>
<tr>
<td>fus6</td>
<td>12.00</td>
<td>0.90</td>
<td>5.92</td>
<td>1.0</td>
<td>0.95 (95)</td>
<td>S</td>
</tr>
<tr>
<td>csu589</td>
<td>12.00</td>
<td>0.90</td>
<td>5.92</td>
<td>1.0</td>
<td>0.95 (95)</td>
<td>S</td>
</tr>
<tr>
<td>bnl5.62a</td>
<td>12.30</td>
<td>0.90</td>
<td>6.07</td>
<td>1.0</td>
<td>0.95 (95)</td>
<td>S</td>
</tr>
<tr>
<td>knox1</td>
<td>14.00</td>
<td>0.89</td>
<td>6.90</td>
<td>1.2</td>
<td>0.94 (94)</td>
<td>S</td>
</tr>
<tr>
<td>umc164c</td>
<td>15.00</td>
<td>0.88</td>
<td>7.40</td>
<td>1.2</td>
<td>0.94 (94)</td>
<td>S</td>
</tr>
<tr>
<td>csu680a</td>
<td>15.90</td>
<td>0.87</td>
<td>7.84</td>
<td>1.4</td>
<td>0.94 (94)</td>
<td>S</td>
</tr>
</tbody>
</table>