Bioinformatics Consortium of Taiwan

http://bct.binfo.org.tw

National Core Facility Program for Biotechnology- Bioinformatics Consortium of Taiwan

PST: Analysis of Human Meiotic Recombination Events with A Parent-Sibling Tracing Approach

Abstract

Meiotic recombination ensures that each child inherits distinct genetic materials from each parent, but the distribution of crossovers along meiotic chromosomes remains difficult to identify. In this study, we developed a parent-sibling tracing (PST) approach from previously reported methods to identify meiotic crossover sites of GEO GSE6754 data set. This approach requires only the single nucleotide polymorphism (SNP) data of the pedigrees of both parents and at least two of children. Analysis of a GEO GSE6754 data set containing 2,145 maternal and paternal meiotic events revealed that the pattern and distribution of paternal and maternal recombination sites vary along the chromosomes. The advantages of the PST approach include the ability to use only two-generation pedigrees with two siblings and the ability to perform gender-specific analyses of repetitive elements and tandem repeat sequences.

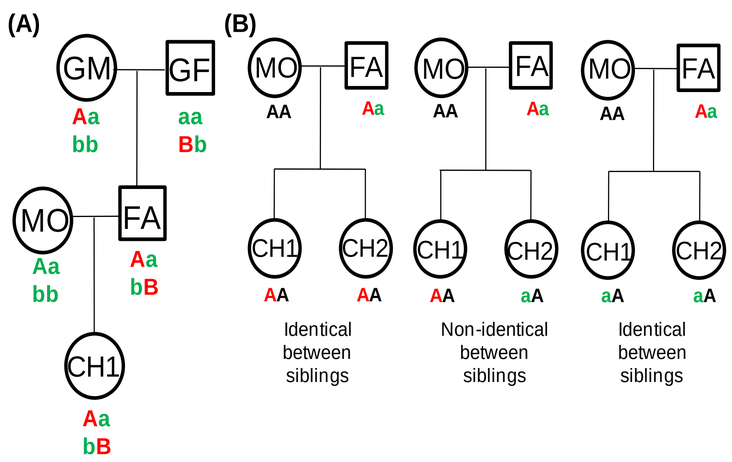

Figure 1. Different types of pedigrees are required for determining meiotic recombination sites by various methods. (A) Three-generation pedigrees are required for the identity by descent (IBD) method, and **(B)** complete two-generation pedigree for the PST method. ***Abbreviations***: GF, grandfather; GM, grandmother; FA, father; MO, mother; CH1 and CH2, child 1 and child 2.

Background

Meiotic recombination is important for generating genetic diversity. Meiotic recombination occurs between homologous chromosomes during chiasmata formation, a process that is required for normal chromosomal segregation during meiosis. Based on the distribution of SNPs in both parents and multiple siblings, meiotic cross sites in human chromosomes can be identified. In this study, we have used a parent-sibling tracing (PST) approach, which was derived from two previous reports (1,2) to analyze the recombination event of GEO GSE6754 data set .

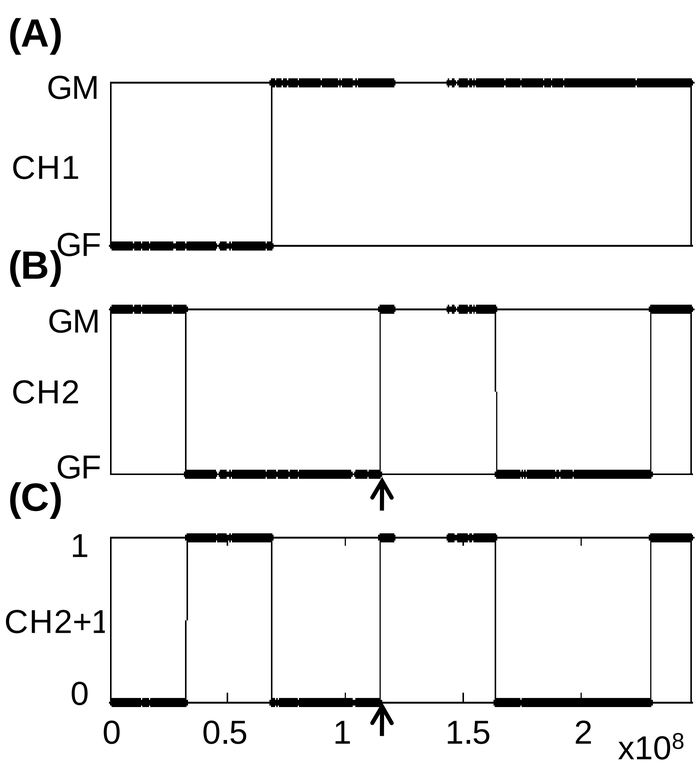

Figure 2. The paternal recombination site on chromosome 1 of child 1 and 2 (CH1 and CH2, defined in Figure 1) in the GMRCL dataset were defined using the identity by descent (IBD) (A, B) and PST (C) methods. The grandmother and grandfather origin of paternal recombination is indicated as GM and GF, respectively. Children with identical or not identical origin are indicated as 1 and 0, respectively.

Results

Comparison of two methods of detecting meiotic recombination sites

We used the GMRCL dataset of 900K SNPs as a reference standard for comparison between the PST approach (Fig. 1B) and IBD method (Fig. 1A). Using chromosome 1 as an example, IBD analysis in both children could define the sites of meiotic recombination for paternal gametes. In child 1 and child 2, we observed 1 and 4 meiotic recombination events on their paternal gametes, respectively (Fig. 2A and 2B). The PST approach (Fig. 2C) detected the recombination sites of the combinatorial results for child 1 and child 2 as determined by IBD (Fig. 2A and 2B).

Analysis of the GEO dataset GSE6754 containing 11,000 SNP markers

The Affymetrix Human Mapping 10K 2.0 Arrays (containing 10K SNPs) were used to map autism susceptibility loci in the GSE6754 dataset. The 3864 arrays (853 families, 1721 parents, 2145 siblings) analyzed using the PST approach, the mean number of maternal recombination events was approximately 1.67-fold higher than that of paternal origin (Fig. 3A).

In order to identify the regions with the highest and the lowest number of recombination events, we scanned the entire human genome and divided the genome into 2,765 bins of 1-Mb each. We then identified the number of recombination sites in each bin separately for female and male meioses (Fig. 3B).

Relation between the recombination site and repetitive elements

We compiled 57 major repetitive element classes that were characterized by RepeatMasker. The correlation coefficients and the corresponding P values for each of the 23 repetitive-elements, CpG island sites, and meiotic recombination sites are summarized in Table 1.

Methods

Identification of meiotic recombination sites

The meiosis recombination sites were exported from the PSTReader, a MATLAB-based program. The source code, example data, and a standalone application can be freely downloaded from: http://www.mcu.edu.tw/department/biotec/en_page/PSTReader/index.htm.

References:

1. Coop et al., Science 2008, 319:1395-1398.
2. Chowdhury et al., PLoS Genet 2009,5:e1000648.

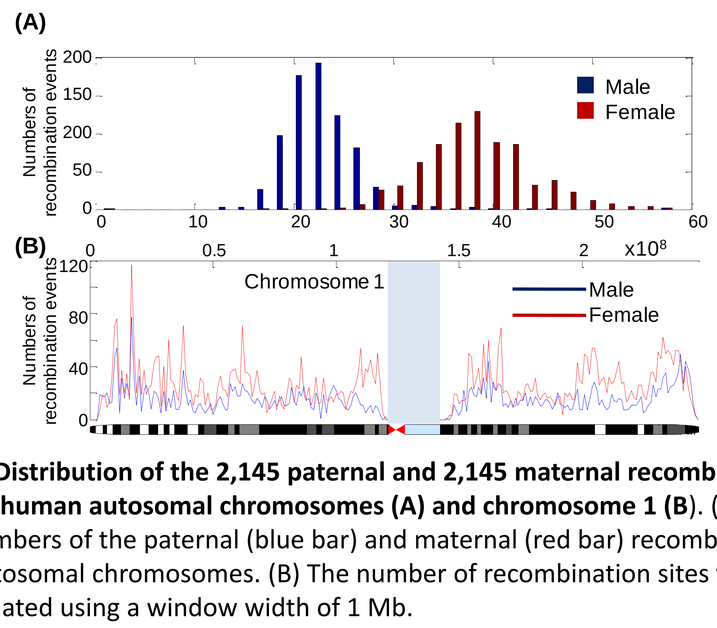

Figure 3. Distribution of the 2,145 paternal and 2,145 maternal recombination events across all human autosomal chromosomes (A) and chromosome 1 (B). (A) The distribution of the numbers of the paternal (blue bar) and maternal (red bar) recombination events across autosomal chromosomes. (B) The number of recombination sites for chromosome 1 was calculated using a window width of 1 Mb.

Table 1. Correlation between the recombination sites and particular repeats

		Paternal		Maternal		Both	
Repeat	**Number**	**Corr.**	**P**	**Corr.**	**P**	**Corr.**	**P**
SINE/MIR	510580	0.22	**1E-16**	0.30	**1E-16**	0.29	**1E-16**
DNA/hAT-Charlie	214901	0.18	**1E-16**	0.31	**1E-16**	0.29	**1E-16**
DNA/hAT	10624	0.16	**3E-13**	0.24	**1E-16**	0.23	**1E-16**
LINE/L2	397294	0.16	**3E-13**	0.22	**1E-16**	0.21	**1E-16**
SINE/Alu	926299	0.10	1E-04	0.22	**1E-16**	0.19	**1E-16**
DNA/hAT-Tip100	26087	0.10	3E-05	0.20	**1E-16**	0.18	**1E-16**
DNA/hAT-Blackjack	17019	0.15	4E-12	0.16	**2E-13**	0.17	**1E-16**
Simple_repeat	343474	0.19	**1E-16**	0.11	**4E-06**	0.16	**8E-14**
Low_complexity	314872	-0.11	**3E-06**	-0.17	**6E-15**	-0.18	**1E-16**
LTR/ERVK	8019	-0.19	**1E-16**	-0.16	**3E-13**	-0.19	**1E-16**
LINE/L1	767428	-0.16	**8E-13**	-0.19	**1E-16**	-0.20	**1E-16**

1. Repeat classes including more than 6000 repeats were considered for the purpose of analyses.
2. Corr.: correlation coefficients between the recombination sites and specific repeats.
3. P values under the null hypothesis of an absence of correlation.
4. The values in bold indicate a P value < 1E5

Welcome to contact with BCT E-mail: bct@ym.edu.tw Tel: 02-28267359 http://bct.binfo.org.tw

NCFPB 生技類核心設施平台維運計畫 National Core Facility Program for Biotechnology