Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 346 (2005) 29–39
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Sex-linked recombination variation and distribution of disease-related genes

Claude Chelala, Charles Auffray*

Genexpress, Functional Genomics and Systems Biology for Health, CNRS UMR 7091-7, rue Guy Moquet, BP8, 94801 Villejuif Cedex, France

Received 5 May 2004; received in revised form 5 August 2004; accepted 14 October 2004
Available online 29 January 2005

Abstract

Analysis of the distribution of recombination along human chromosomes and correlation with sequence features and genes have been previously performed on one genetic map for a given chromosome, limiting therefore their validity and precision. In this paper, we circumvent these issues: (1) by testing the correlation between recombination frequency in sex-specific versions of three genetic maps of chromosome 21 and their content in disease-related loci compared to the distribution of genes along the chromosome, and (2) by reanalysing the previously reported chromosome 22 results (Chelala et al., J. Biol. Syst. 10 (2002) 303–317) with updated version of the sequence and mapping tools.

Recombination hot zones were detected and analysed on each genetic map. Despite local differences, for chromosome 21, recombination hot zones were found relatively enriched in disease-related genes on the *male* genetic maps. This contrasts with the previously described enrichment of the chromosome 22 *female* genetic map hot zones in disease-related loci (Chelala et al., J. Biol. Syst. 10 (2002) 303–317), which was confirmed with the updated data and tools.

Our study demonstrates that the use of different data sets and tools have only a local impact on the distribution of genetic recombination hot zones and provides evidence for gender-specific differences in enrichment in disease-related loci in relation with recombination frequency. Automation of such analyses and extension to the entire human genome will be required in order assess the general character of these observations and to advance in the understanding of genome-wide recombination patterns to help the process of identifying disease-causing genes.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Genetic map; Chromosome 21; Chromosome 22; Recombination hot spot

1. Introduction

Most comparison studies of human genetic and sequence-based physical maps have been performed with genetic maps built using genotyping data from a limited number of CEPH families comprising a relatively small number of meioses and providing therefore only limited information (Broman et al., 1998; Dunham et al., 1999; Hattori et al., 2000; Majewski and Ott, 2000; Yu et al., 2001; Chelala et al., 2002; Tapper et al., 2002). The order of genetic markers was established according to the genomic sequence, the relationship between the genetic and physical distances were analysed, resulting in recombination hot and cold spot maps for each chromosome. Through such analyses, it was reported that regions with increased recombination frequency are correlated with the abundance of repetitive DNA (Majewski and Ott, 2000; Tapper et al., 2002) or enriched in genes (Lynn et al., 2000) and disease-related genes (Chelala et al., 2002).

Recombination along the human genome is not random. It depends on individual gender, chromosome structure and

Abbreviations: cM, centiMorgan; cR, centiRays; Mb, mega (million) base pairs; kb, kilo base pairs.

* Corresponding author. Tel.: +33 1 49 58 34 97; fax: +33 1 49 58 35 09.
E-mail address: auffray@vjf.cnrs.fr (C. Auffray).

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.10.019

sequence composition. It is well known now that sex-average recombination rates, which are the average for males and females combined, do not vary as much as the sex-specific rates calculated for males and females considered separately. This is due to the fact that male and female recombination rates at specific sites often differ substantially. Thus, studying sex-specific maps is preferred to detect regions that are "hot spots" or "cold spots" for one or both genders.

However, the analysis of one genetic map cannot rule out the possibility that the results obtained might be due to problems associated with the genetic and/or physical data used, such as limited information, weak resolution and order errors (Matise et al., 2002). It is therefore important to address the question of whether the recombination hot zones and/or correlations previously reported can change or disappear once a different set of genetic data is analysed. Although many studies have been performed to analyse the distribution of chromosomal hot and cold spots of recombination, to date, there are no published results formally confirming the original distribution and correlations while considering another set of genetic data.

In order to test the validity of the reported observations, we analysed three genetic maps of chromosome 21 against National Center for Biotechnology Information, NCBI build 31 of the human genome sequence. We report that recombination hot zones are relatively enriched in disease-related genes on the male genetic maps. We also updated the work we previously performed on chromosome 22 (Chelala et al., 2002) in order to test if there is any interference of the sequence updates and/or the method used to built genetic maps on the results obtained. We confirm that chromosome 22 female genetic map hot zones are enriched in disease-related loci.

2. Data and methods

2.1. Data

2.1.1. Genetic data

Three sets of markers were used to build three genetic maps for chromosome 21. Set 1 is derived from the NCBI data of 48 markers from the Généthon map (Dib et al., 1996), whose order is consistent with that in the Marshfield map (Broman et al., 1998) and is placed on the sequence map. Set 2 contain 78 markers from the Marshmed genetic map that are placed on the genome sequence (Broman et al., 1998; Yu et al., 2001). Set 3 contain 51 markers, from the first and second sets combined, having genotype data from the deCODE mapping project (Kong et al., 2002).

The chromosome 22 map contains 68 markers derived from the Marshfield map (Broman et al., 1998; Yu et al., 2001). It extends over 72 and 42.3 centiMorgan (cM) on the female and male genetic maps, respectively, covering 31.19 mega (million) base pairs (Mb) of the sequence.

2.1.2. Physical data

The sequence coordinates for all markers were collected from the NCBI build 31 sequence data. The chromosome 21 sequence consists of five contiguous segments covering 33.92 Mb spaced by four gaps estimated to cover 3.12 Mb. The chromosome 22 sequence consists of 11 contigs of 33.82 Mb separated by 10 gaps of 9.27 Mb.

2.1.3. Transcript data

The 'Genes On Sequence' map from the Mapviewer at NCBI was used, representing 300 genes that have been annotated on the chromosome 21 genomic contigs and 697 genes on the chromosome 22 genomic contigs.

2.1.4. Morbid data

Online Mendelian Inheritance in Man, OMIM and LocusLink were searched for chromosome 21-related phenotypes. For 23 phenotypes, the corresponding 19 loci sequence coordinates were collected from the transcript map of chromosome 21. We considered 41 chromosome 22 loci implicated in 32 diseases and covered by the genetic and sequence maps.

2.1.5. Cytogenetic data

The cytoband information was collected from the NCBI mapping information contained in the ISCN800_abc file. It includes the chromosomal distance in megabases and the corresponding cytoband (based on ISCN 800 level band resolution size measurements; Mitelman, 1995). More detailed description as well as updated information of the physical, transcript, morbid and cytogenetic data can be found at the NCBI web site.

2.2. Methods

2.2.1. Construction of genetic maps

Using the markers order on the sequence, the genetic map distances (in cM) of polymorphic markers for the first two sets of genetic markers were determined using the MAP-O-MAT Web-based server that automates the use of the CRI-MAP program (Lander and Green, 1987; Matise and Gitlin, 1999). MAP-O-MAT uses the CEPH genotype database version 8.2, comprising 178 chromosome 21 markers scored in the CEPH reference pedigrees. With the sequence map order for markers, MAP-O-MAT allows to perform a FIXED analysis to estimate the intermarker recombination fractions and map distances, and a FLIPS analysis to evaluate local support for order. CRIMAP was used directly, rather than through MAP-O-MAT, to analyse the third set of markers against the deCODE genotype data. In general, when a lod score is greater than 3, it provides substantial statistical evidence in favor of the markers order. When it is negative, it provides evidence in favor of the

Download English Version:

https://daneshyari.com/en/article/9127146

Download Persian Version:

https://daneshyari.com/article/9127146

Daneshyari.com