Short Communication

tSNP-based identification of allelic loss of gene expression in a patient with a balanced chromosomal rearrangement

Gregory F. Guzauskas, Kennedy Ukadike, Lynn Rimsky, Anand K. Srivastava *

J.C. Self Research Institute of Human Genetics, Greenwood Genetic Center, 113 Gregor Mendel Circle, Greenwood, SC 29646, USA

Received 23 August 2006; accepted 12 December 2006
Available online 22 January 2007

Abstract

Identification of genes affected by disease-associated rare chromosomal rearrangements has led to the cloning of several disease genes. Here we have used a simple approach involving allele-specific RT-PCR-based detection of gene expression to identify a gene affected by a balanced autosome;autosome translocation. We identified a transcribed SNP (tSNP), c.68G→A, present in a novel untranslated exon of the CLDN14 gene in a male patient with mental retardation who had a balanced t(13;21) chromosomal translocation. We determined an allelic loss of expression of the CLDN14 gene isoform at the 21q22.1 chromosomal breakpoint. Although additional work is necessary to explore a possible function of the novel CLDN14 isoform in brain development and function and the potential pathogenic consequences of its disruption in this patient, the result clearly demonstrates the utility of a tSNP-based detection of allelic loss of gene expression in studies involving chromosomal rearrangements.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Chromosome translocation; tSNP; RT-PCR; Allele-specific expression

The identification of disease genes through positional cloning approaches has proven to be one of the most powerful tools in unraveling gene function. Autosomal dominant diseases are normally caused by a defect in one of two copies of an autosomal gene. Such defects lead to either haploinsufficiency or a gain in gene function. Identification of causative autosomal genes in diseases of a heterogeneous nature such as mental retardation and autism has proven to be very difficult due to the lack of large families for linkage analysis [1]. Furthermore, phenotypic variability among small families, often consisting of single sibships, precludes the pooling of families for linkage analysis. Thus, progress in the identification of autosomal mental retardation (MR) genes has been very limited [2–4].

Studies involving disease-associated rare chromosomal rearrangements, such as balanced translocations, inversions, and duplications, have been very productive in identifying the causative genes [5,6]. However, mapping of the chromosomal breakpoints is a labor-intensive and multistep process and requires delineation of the critical chromosomal breakpoint region and identification of candidate genes from the breakpoint critical region [5–8]. In most of these cases, the causative genes have been found to be physically disrupted by the chromosomal breaks. However, in a number of cases, a long-range effect of the chromosomal break on genes located at a distance has also been identified [9]. Although productive, there are limitations in the use of traditional approaches for detecting genes disrupted or affected by a chromosomal break.

Real-time RT-PCR expression analysis to determine expression variations in genes affected by chromosomal rearrangements has also been utilized in a limited number of studies [10]. However, this has not been a method of choice due to the wide range of observed expression variability of reference and experimental genes among test or control samples [11,12]. Allele-specific gene expression has been extensively utilized in studies of gene silencing related to X inactivation and imprinting studies [13–15]. Here we have used a transcribed SNP (tSNP)-based RT-PCR approach and show its utility in the identification of a gene affected by a chromosomal breakpoint in a 19-year-old patient with a balanced t(13;21) (q22;q22.1) translocation, mild MR (IQ 69), macrocephaly (OFC >98th percentile), and aggressive behavior.

* Corresponding author. Fax: +1 864 388 1808.
E-mail address: anand@ggc.org (A.K. Srivastava).

0888-7543/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
To identify the potential candidate gene(s) for the observed clinical features in the patient, we mapped the 21q22.1 breakpoint to an approximate 50-kb genomic region flanked by two previously characterized genes, CLDN14 and SIM2 (cen–CLDN14–21q breakpoint–SIM2–ter) using fluorescence in situ hybridization (data not shown). It has been suggested that an extra copy of SIM2 may be important for the pathogenesis of Down syndrome, especially mental retardation (MIM 600892) [16]. However, a role for the Down syndrome critical region genes in specific Down syndrome phenotypes has been debated [17]. The SIM2 gene mapped approximately 200 kb telomeric to the breakpoint and was apparently not affected by the translocation breakpoint.

The CLDN14 gene comprises three exons spanning (tel→cen) about 19.4 kb of genomic DNA. The first two exons of the CLDN14 gene encode a 5′ untranslated region; exon 3 contains the complete 720-bp open reading frame, encoding a 239-amino-acid protein of the claudin gene family. An isoform of the CLDN14 gene has been previously shown to be mutated in some cases of nonsyndromic deafness (DFNB29) (MIM 605608) [18–20]. No deafness was noted in the present patient.

To determine if one of the two genes or a new previously unidentified expressed sequence is affected by the 21q breakpoint in the patient, we performed computer-assisted sequence analysis of the breakpoint critical region. A limited number of EST sequences were identified within the breakpoint interval. Further EST sequence analysis and analysis of the corresponding mouse sequences led to identification of three novel 5′ untranslated exons and two new transcript isoforms of the CLDN14 gene (variants 3 and 4; Fig. 1). These findings also complemented recently identified novel isoforms of the CLDN14 gene [20]. Altogether, identification of novel CLDN14 isoforms extended the 5′ end of the CLDN14 gene further within the 21q breakpoint region and suggested a likely physical disruption of the CLDN14 gene (Fig. 1) by the 21q22.1 breakpoint. We hypothesized that the chromosomal break at 21q22.1 should lead to a loss of expression of a copy of a CLDN14 gene isoform located on the derivative chromosome 21.

To check for an effect of the 21q chromosomal break on the expression of a novel CLDN14 isoform, we performed quantitative RT-PCR analysis. However, results were inconclusive. Therefore, we devised an alternative approach to look for the expression of the CLDN14 gene using allele-specific RT-PCR. First, we identified a heterozygous nucleotide variation (c.68G→A) in the novel untranslated exon 1a of the CLDN14 gene in the patient (Figs. 1 and 2A, top). We used this tSNP to examine specifically the expression of one of the two isoforms of the CLDN14 gene (Fig. 1, variant 3) that was predicted to be disrupted by the translocation breakpoint. We performed allele-specific RT-PCR analysis on total RNA extracted from the patient’s EBV-transformed lymphoblasts as described previously [21]. First-strand synthesis was performed using the SuperScript first-strand synthesis system for RT-PCR (Invitrogen). A cDNA amplicon was amplified using primers specific for CLDN14 exon 1a (5′-TGCTGGGCTGCTGGAAGTCCT-3′) and exon 1b (5′-AAGCGCCATGAGCTGGTGTG-3′) in a reaction volume of 25 μl that contained 10 μM each primer,

**Fig. 1.** A schematic representation of the 21q22.1 chromosomal breakpoint region in the patient with the t(13;21) translocation. Arrows indicate the direction of the transcription. The entire open reading frame of the CLDN14 gene is present in exon 3. The novel alternative exons (1a, 1b, and 1c) and transcript variants/isoforms of the CLDN14 gene are shown. The respective GenBank accession numbers of the variants are in parentheses. Three tSNPs, c.68G/A in alternative exon 1a (bottom strand) and c.11C/A and c.63G/A in coding exon 3 (top strand), are indicated.
250 μM dNTPs, 1× PCR buffer, and 1 unit of Taq polymerase (Sigma) with 0.02 μM TaqStart antibody (Clontech). The RT-PCR conditions included an initial denaturation at 95°C for 300 s, followed by 40 cycles of 95°C for 30 s, 69°C for 30 s, and 72°C for 30 s. The PCR product (140 bp) was gel purified and both strands were sequenced using the DYEEnamic ET Dye Terminator Cycle Sequencing Kit (GE Healthcare) and the automated MegaBACE 1000 DNA analysis system (GE Healthcare). Sequences were analyzed using the DNASTAR program (Madison, WI, USA). Sequence analysis revealed that the patient’s cDNA carried only a “G” allele at nucleotide 68 (Fig. 2A). Taken together, the results indicate that only one copy of CLDN14 was transcribed and suggested a loss of expression of the second copy of CLDN14 located on the derivative chromosome 21.

To exclude a possibility of an imprinting effect on the CLDN14 gene expression that may have caused the monoallelic expression of the CLDN14 gene in the patient, we analyzed both genomic DNA and cDNA from two control individuals who were also heterozygous for the c.68G/A variation (Fig. 2B). Taken together, the results indicate that only one copy of CLDN14 was transcribed and suggested a loss of expression of the second copy of CLDN14 located on the derivative chromosome 21.

As expected, the results were further confirmed by amplifying the cDNA products using a primer from exon 1b and exon 3 (variant 3). Furthermore, two additional tSNPs, c.11C/T and c.63G/A, located in coding exon 3 were also examined in the patient and in two controls (Fig. 1 and data not shown).

The functional significance of the CLDN14 isoforms containing untranslated exons remains to be determined. A role for the smaller CLDN14 isoform containing exon 1, exon 2, and exon 3 in autosomal recessive deafness has been reported [18–20]. The patient studied here reportedly has no deafness. Although a likely role for one of these isoforms in human brain function is predicted, the present data are not sufficient to draw any such conclusion and would require additional functional studies. Nonetheless, the method described here was clearly able to show the utility of allele-specific RT-PCR in determining the effect of the translocation breakpoint on the transcription of one copy of the CLDN14 gene in the patient.

Polymorphic genomic markers and single-nucleotide polymorphic alterations have been extensively used in genotype analysis using genomic DNA. SNPs as well as other polymorphic markers have also been used extensively to infer genomic deletions, in studies involving uniparental disomy and loss of heterozygosity, and in imprinting studies [13,22,23]. The approach described here utilizes the presence of a heterozygous nucleotide alteration in transcribed regions to identify genes affected by the translocation breakpoints.

In most autosomal dominant cases, a chromosomal rearrangement would likely result in the loss of expression of one copy of a gene either by physical disruption of the gene or due to its presence in close vicinity of the breakpoint. However, in the latter case, an overexpression of the gene due to physical separation of a regulatory element by the breakpoint cannot be excluded. Detection of such long-range effects of the autosomal breakpoints on a gene(s) located at some distance from the breakpoint remains problematic [24]. Utility of real-time expression analysis using quantitative RT-PCR has also been very limited [10]. Thus, the tSNP-based RT-PCR approach described here provides a quick and efficient screening method to analyze a large number of genes from the breakpoint critical regions or regions flanking chromosomal breakpoints. A similar approach has been
successfully applied in some cases with unbalanced X; autosome translocations to examine spreading of X inactivation in the translocated segment of the autosomes [25]. The approach can be easily adapted to determine quantitative expression difference due to difference in copy number of the gene utilizing any tSNP [26,27] via information available through the SNP database (http://www.ncbi.nlm.nih.gov). It is important to note that this approach is increasingly feasible because there are now over 5,000,000 SNPs available for potential use for screening of genes and isoforms, and the annotation of the genome map is very informative. The candidate genes from the breakpoint regions could be easily monitored for the presence of informative heterozygous tSNPs.

In the present study, we tested and provide a simple and efficient approach for identifying genes affected by chromosomal breakpoints. We have been applying this tSNP-based approach to screen and identify genes affected by the translocation breakpoint in other cases with chromosomal rearrangements. We have successfully analyzed a candidate gene in a female patient with a t(3;9) translocation. Furthermore, we were able to screen genes present at a distance and flanking the breakpoints in a patient with a t(2;11) translocation and determined a normal biallelic expression. This suggests that the approach can also be efficiently used to screen a number of genes from the chromosomal breakpoint critical regions to determine a long-range effect of the chromosomal break on the gene expression when the breakpoint does not necessarily disrupt a gene.

Acknowledgments

We thank K.B. Clarkson and C. Skinner for assistance in obtaining the patient’s sample, D. Schultz for assistance in sequencing, Ada King for technical assistance, and C.E. Schwartz for helpful discussion. This work was supported by a grant from the National Institutes of Child Health and Development (ROI-HD39331).

References