Deep Insight Section

LDI-PCR in Cancer Translocation Mapping

Björn Schneider, Hans G Drexler, Roderick AF MacLeod

DSMZ, German Collection of Microorganisms and Cell Cultures, Department of Human and Animal Cell Cultures, Inhoffenstr. 7b, 38124 Braunschweig, Germany (BS, HGD, RAFM)

Published in Atlas Database: April 2010
DOI: 10.4267/2042/44951

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 2.0 France Licence.
© 2011 Atlas of Genetics and Cytogenetics in Oncology and Haematology

Abstract

Identification of genes in oncogenic chromosome translocations by Fluorescence In Situ Hybridization (FISH) screening using genomic tilepath clones, is often laborious, notably if the region of interest is gene-dense. Other molecular methods for partner identification also suffer limitations; for instance, genomic PCR screening requires prior knowledge of both sets of breakpoints, while Rapid Amplification of cDNA Ends (RACE) is limited to translocations causing mRNA fusion and delivers no breakpoint data. With Long Distance Inverse (LDI)-PCR, however, it is possible to identify unknown translocation partners and to map breakpoints at the base-pair level. Applying LDI-PCR merely requires approximate sequence information on one partner, rendering it ideal for use in combination with FISH to extend and refine cytogenetic breakpoint data.

Introduction

Recurrent chromosomal rearrangements characterize many different types of cancer. Specific cytogenetic translocations are key events, widely considered to be diagnostically and prognostically significant in leukemia and lymphoma, and increasingly so in solid tumors (Mitelman et al., 2007). Hitherto, most cancer genes have been identified following analysis of recurrent chromosome translocations (Futreal et al., 2004). The pathological significance and usefulness of such rearrangements depend on two key features: a) whether rearrangements display distinct patterns of recurrence within specific tumors, e.g. t(8;14)(q24;q32) which is restricted to B-cell neoplasia; and b) how clustered are the chromosomal breakpoints therein. The significance accorded to breakpoint data depends on their ascertainment precision, from megabase- and kilobase-down to single base-pair levels, when ascertained by classical cytogenetics, fluorescence in situ hybridization (FISH), and sequence-based methods, respectively.

Chromosome translocations fall into three broad categories. The first causes the physical fusion of the two mRNAs expressed by the participating genes, thus creating novel fusion proteins translated from exons emanating from both genes, e.g. BCR (at chromosome-22-[band]-q11) with ABL1 (at 9q34) fused by t(9;22)(q34;q11) in chronic myeloid leukemia (CML) and in some cases of acute lymphoblastic leukemias (ALL) (Turhan, 2008a,b). The second category also fuses mRNA from genes at separate loci but, in this case, serving to deregulate a developmentally silenced partner by exchanging promoters with more active partners, e.g. BCL6 (at 3q27) which is activated by translocations with any one of many partners, chiefly in diffuse large B-cell lymphoma (DLBCL) (Knezevich, 2007).

The third class of chromosome translocation again results in the activation of the normally silent partner, this time by juxtaposition with another constitutively active partner without mRNA fusion, e.g. the neighboring homeobox genes, TLX3 (at 5q35.1) and NKX2-5 (at 5q35.2). According to the proximity of the breakpoint involved, either (but not both) genes may be activated in T-cell ALL by the recurrent t(5;14)(q35;q32.2) by which these are juxtaposed with regulatory regions from BCL11B (at 14q32.2) to stimulate transcription (Bernard et al., 2001; MacLeod et al., 2003; Nagel et al., 2007).

Some "promiscuous" genes engage with multiple partners: notably MLL with 64 known partners (Meyer...
et al., 2009b), BCL6 with 28 (Knezevich, 2007), RUNX1 with 39 (Huret and Senon, 2003), NUP98 with 29 (Kearney, 2002), and the IgH-locus with 40 (Lefranc, 2003).

Although, promiscuity may reflect the dependence of tumors on the inappropriate oncogene expression without overly caring how deregulation is accomplished, the role of the partner genes has come under renewed scrutiny. Choice of partner gene may not only reveal in which types of precancerous cells primary oncogenic rearrangements occur, but also by looking for conserved DNA or protein motifs, yield clues to the mechanisms underlying their formation or their functional contribution to neoplasia. In addition, the roles of biologically important genes, e.g. BCL11B, a key regulator of both differentiation and survival during thymocyte development, are often first rendered visible by their participation in cancer rearrangements (MacLeod et al., 2003).

Both the identities of each partner gene and their precise breakpoints at the DNA base-pair level, may be useful not only to characterize potential fusion genes/products but also to ascertain whether additional non-protein-coding genomic entities, such as chromosomal fragile sites (Schneider et al., 2008), microRNA loci, putative genes, unspliced "expressed sequence tags", or regulatory non-coding regions may be involved.

Clues to the biological mechanisms generating chromosome rearrangements are given by breakpoint sequences, including T- and B-cell receptor gene (VDJ) rearrangement, Alu-mediated recombination, non-homologous end joining, etc. VDJ genes are flanked by recombination signal sequences composed of heptamers followed, in turn, by a spacer containing either 12 or 23 unconserved nucleotides and a conserved nonamer. Spacers of 12 nucleotides undergo physiological recombination with those containing 23 in order to obey the so-called "12/23 rule". The presence of these or related sequences has been reported in connection with cancer translocations to reveal how physiologic processes may be abused to cause genomic rearrangements (Gu et al., 1992).

Genomic fusion sequences may also be used to aid cell line authentication - an omnipresent problem confronting cell culturists, given that an unexpectedly (and unacceptably) high percentage of new cell lines has been misidentified, or cross-contaminated by older cell lines (MacLeod et al., 1999). While mRNA fusion sequences are constrained by splicing, their genomic equivalents allow sufficient variation to provide "fingerprints" unique to individual cell lines to serve as potential identifiers. For analyzing patient tumors, knowledge of the exact fusion sequence allows design of patient-specific quantitative (q)PCR used for monitoring minimal residual disease with high sensitivity (Burmeister et al., 2006), to follow up therapeutic responses thus enabling early detection of relapse.

**Figure 1:** Amplifying Genomic Fusions of Unknown Sequence. The schema summarizes how the genomic DNA is first restricted, then religated to the circular template, and how the resulting amplicon should appear. Note unknown region (red) flanked by known sequences (black). R: restriction site, BP: breakpoint; arrows: forward (FW) and reverse (REV) primers.
Cancer gene promiscuity enables oncogenic chromosome rearrangements, "smoking guns" of cancer genes, to be distinguished from random changes. FISH is initially used to confirm rearrangement of a contextually appropriate oncogene residing at the locus in question. Hence, a breakpoint at 9q34 might throw suspicion onto NOTCH1 in a T-cell-, ABL1 in myeloid- neoplasia, and NUP214 in either entity. While such an approach is less helpful among solid tumors where oncogene rearrangements are less informative, at this locus TSC1 might be deemed a candidate in tuberous sclerosis cells. Even when the index breakpoint is precisely known, determination of its partner by FISH requires time-consuming and laborious procedures for those not afforded immediate blanket tilepath-clone coverage with which to quarter the region of interest.

PCR screening with hit-lists of known and potential partner genes is quite as laborious as FISH, and is liable to miss unknown translocation partners, or those with breakpoints lying outside their respective cluster regions. When there are grounds to suspect transcriptional fusion (as among partners of genes prone to this type of gene rearrangement, e.g. ABL1, ETV6, NUP98, etc.) a mRNA-based method, rapid amplification of cDNA ends (RACE), may be used to detect novel fusion partners (Frohman et al., 1988). A drawback of RACE is its inability to supply genomic breakpoint data, and the risk of overlooking some splice variants. Hence, the technique of choice for identifying unknown partner genes and their breakpoints should not require prior knowledge of the partner gene, yet provide breakpoint data at the DNA base pair level.

Long Distance Inverse (LDI)-PCR satisfies these needs. LDI-PCR was developed from the earlier inverse-PCR (Ochman et al., 1988) to allow the amplification of large DNA fragments comprised of known and unknown sequences (Willis et al., 1997) using re-ligated circular restriction fragments as templates. Primers are set in opposition within the known sequence. The unknown sequence is flanked on both sides by known sequences following re-ligation in the resultant amplicon (Fig. 1). When a restriction fragment length polymorphism (RFLP) distinguishing the wild type and derivative alleles is generated by the genomic alteration, the two resulting amplicons should be separable by gel electrophoresis, enabling their respective sequences to be compared (Fig. 2). Sequencing with one of the PCR-primers directed towards the restriction site allows immediate identification of the partner gene. Sequencing in the other direction allows precise mapping of the breakpoint.

Figure 2: How to Interpret LDI-PCR Gels. Left figure shows the wild type configuration where twin circular templates identical in size would yield a single band by agarose gel electrophoresis. Translocation bearing cells (right figure) yield both wild type and derivative templates, differing in size and detectible as two bands on the gel. The derivative band is indicated by an arrow. Known regions are outlined in black, unknown in red. R: Restriction site, REV: reverse primer, FW: forward primer.
While easier to perform for genes with well defined, short breakpoint cluster regions, LDI-PCR may be applied to any gene or region involved in a translocation and has, therefore, been applied to a wide variety of translocations involving both frequently rearranged promiscuous oncogenes, but also as single events. Table 1 gives an overview of genes analyzed by LDI-PCR according to literature databases and the respective references.

Limitations are set by the performance of the DNA polymerase since lengthier fragments may resist amplification, and by the placement of the RFLP, as fragments similar in size cannot be readily distinguished by gel electrophoresis. If primary patient tumor material is analyzed, it should be noted that the samples used for analysis not only contain tumor material, but also normal bystander cells devoid of tumor rearrangement. Detection attempted at lower tumor infiltration rates risk false negative results.

In contrast to other PCR methods suitable for detection of unknown fusion sequences, such as panhandle PCR (Megenigal et al., 2000) or analogous techniques requiring adaptor ligations (reviewed in Tonooka and Fujishima, 2009), LDI-PCR is independent of any additional adaptors or anchors which have to be ligated to the restricted fragments, reducing the number of steps required, while remaining sufficiently flexible to allow a wide choice of restriction enzymes.

In the future, translocation analysis by next generation sequencing should overcome these limitations and suitable algorithms have been developed to recognize novel derivative breakpoint-flanking sequences and thereby identify novel cancer translocations and other synonymous rearrangements, including a subset of fusogenic microdeletions (Campbell et al., 2008).

**Methology**

In principle, LDI-PCR utilizes digested and re-ligated circular templates, which are of different sizes, due to RFLP caused by genomic rearrangements. This size difference renders the amplicons separable by gel electrophoresis (Fig. 2).

For a successful analysis, the LDI-PCR has to be designed carefully. The sequence covering the genomic region of interest should be selected from a genome browser (ENSEMBL, UCSC, NCBI) and then pasted into the query box of a restriction map generator (BioEdit, multiple online tools: SMS, RestrictionMapper). Restriction enzymes should be chosen to yield fragments in a size range of 2-5 kb. If using a double-digest strategy with enzymes producing sticky ends, these ends must be compatible. Ensure that both enzymes perform well in the same buffer and at the same temperature. Primer pairs have to be designed in such a way that one primer is directed towards the restriction site, the other one in the opposite direction (see Figures). The sequence lying between the primer tails is not subject to amplification, so the gap should not be excessive, ideally 30-50 bp. A breakpoint lying therein cannot be detected unless another primer pair, e.g. at the other end of the restriction fragment is used. For longer fragments (greater than 5 kb, say) use of a primer set consisting of one forward and multiple reverse primers (or vice versa) can be helpful. The oligonucleotides should be ~30 bp with a Tm ~65°C and a GC-content of 40-60%.

For LDI-PCR template preparation high quality genomic DNA should be used, meaning high purity (260/280 1.8-2.0 and 260/230 > 2) and high integrity without degradation. One microgram of DNA is then digested with 30-50 U of each restriction enzyme in the presence of the appropriate digestion buffer in a total volume of 100 μl for 3-4 h at the temperature suitable for the chosen enzymes (mostly 37°C), followed by heat inactivation (where applicable) and purification, preferably with a column based purification kit. Phenol / chloroform purification followed by precipitation may also be performed, but residual phenol can disturb downstream processes. To form the circular templates, the restriction fragments are then religated with 5 U T4 ligase overnight at 4-8°C in a total volume of 80 μL, terminated by heat inactivation. These conditions favor the desired self-ligation.

The PCR is performed best using a PCR kit suitable for amplification of long templates and using 5 μl (62.5 ng) of the digested and re-ligated DNA. The PCR products are analyzed by gel electrophoresis. Discrepant bands not corresponding to the calculated amplicon size may represent amplicons of translocated fragments. These are excised from the gel, purified and subjected to sequence analysis and, unless artefacts, may reveal the translocation partner and the exact breakpoint of the rearrangement subject to analysis.
# LDI-PCR in Cancer Translocation Mapping

**Schneider B, et al.**

**Atlas Genet Cytogenet Oncol Haematol.** 2011; 15(1)

<table>
<thead>
<tr>
<th>Gene</th>
<th>No. of Partners</th>
<th>References describing LDI-PCR analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTXR1</td>
<td>1 (Oberthuer et al., 2005)</td>
<td>Oberthuer et al., 2005</td>
</tr>
<tr>
<td>API2</td>
<td>1 (Mathijs and Marynen, 2001)</td>
<td>Dierlamm et al., 1999</td>
</tr>
<tr>
<td>BCL6</td>
<td>28 (Knezevich, 2007)</td>
<td>Akasaka H et al., 2000; Akasaka T et al., 2000; Kurata et al., 2002; Akasaka et al., 2003; Chen et al., 2003; Montesinos-Rongen et al., 2003; Chen et al., 2006; Schneider et al., 2008</td>
</tr>
<tr>
<td>BRD4</td>
<td>1 (Collin, 2007)</td>
<td>Haruki et al., 2005</td>
</tr>
<tr>
<td>CALL</td>
<td>1 (Frints et al., 2003)</td>
<td>Frints et al., 2003</td>
</tr>
<tr>
<td>E2A</td>
<td>5 (Huret, 1997)</td>
<td>Wiemels et al., 2002a</td>
</tr>
<tr>
<td>ETV6</td>
<td>28 (Knezevich, 2005)</td>
<td>Wiemels et al., 1999a; Wiemels et al., 1999b; Wiemels and Greaves, 1999; Wiemels et al., 2008</td>
</tr>
<tr>
<td>IGH</td>
<td>40 (Lefranc, 2003)</td>
<td>Willis et al., 1997; Willis et al., 1998; Nardini et al., 2000; Satterwhite et al., 2001; Sonoki et al., 2001; Bichi et al., 2002; Sanchez-Izquierdo et al., 2003; Sonoki et al., 2004; Akasaka et al., 2007; Lenz et al., 2007; Souabni et al., 2007; d’Amore et al., 2008; Ishizaki et al., 2008; Russell et al., 2008; Vieira et al., 2008; Vinatzer et al., 2008; Nagel et al., 2009; Russell et al., 2009; Yin et al., 2009; Hu et al., 2010</td>
</tr>
<tr>
<td>let-7a-2, miR-100</td>
<td>1 (Bousquet, 2008)</td>
<td>Bousquet et al., 2008</td>
</tr>
<tr>
<td>MLH1</td>
<td>1 (Meyer et al., 2009b)</td>
<td>Meyer et al., 2009a</td>
</tr>
<tr>
<td>MLL</td>
<td>64 (Meyer et al., 2009b)</td>
<td>Blanco et al., 2001; Meyer et al., 2005; Teuffel et al., 2005; Attarbaschi et al., 2006; Burmeister et al., 2006; Matsuda et al., 2006; Meyer et al., 2006a; Meyer et al., 2006b; Strehl et al., 2006; Burmeister et al., 2008; Balgobind et al., 2009; Bueno et al., 2009; Burmeister et al., 2009; Matsuda et al., 2009; Meyer and Marschalek, 2009; Cósé et al., 2010; De Braekeleer et al., 2010; Lee et al., 2010</td>
</tr>
<tr>
<td>NOTCH1</td>
<td>2 (Suzuki et al., 2009)</td>
<td>Suzuki et al., 2009</td>
</tr>
<tr>
<td>PAX5</td>
<td>6 (Strehl, 2005)</td>
<td>An et al., 2008; An et al., 2009</td>
</tr>
<tr>
<td>PDGFRA</td>
<td>6 (Dessen, 2009)</td>
<td>Cools et al., 2003</td>
</tr>
<tr>
<td>PDGFRB</td>
<td>20 (Vizmanos, 2005)</td>
<td>Walz et al., 2007; Walz et al., 2009</td>
</tr>
<tr>
<td>RUNX1 / AML1</td>
<td>39 (Huret and Senon, 2003)</td>
<td>Xiao et al., 2001; Wiemels et al., 2002b</td>
</tr>
</tbody>
</table>

**Table 1:** Genes involved in translocations analyzed by LDI-PCR, numbers of yet known translocation partner genes and references wherein the analysis of the particular gene is described.

**References**


MacLeod RA, Dirks WG, Matsuo Y, Kaufmann M, Milich H, Drexler HG. Widespread intraspecies cross-contamination of human tumor cell lines arising at source. Int J Cancer. 1999 Nov 12;83(4):555-63


MacLeod RA, Nagel S, Kaufmann M, Janssen JW, Drexler HG. Activation of HOX11L2 by juxtaposition with 3′-BCL11B in acute lymphoblastic leukemia cell line (HPB-ALL) with t(5;14)(q35;q32.2). Genes Chromosomes Cancer. 2003 May;37(1):84-91


Attarbaschi A, Nagel S, Kaufmann M, Janssen JW, Drexler HG. Activation of HOX11L2 by juxtaposition with 3′-BCL11B in acute lymphoblastic leukemia cell line (HPB-ALL) with t(5;14)(q35;q32.2). Genes Chromosomes Cancer. 2003 May;37(1):84-91


Nagel S, Scherr M, Kel A, Hornischer K, Crawford GE, Kaufmann M, Meyer C, Drexler HG, MacLeod RA. Activation of TLX3 and XKX2-5 in (5;14)(q35;q32) T-cell acute lymphoblastic leukemia by remote 3′-BCL11B enhancers and
recombinome of acute leukemias. Leukemia. 2009 Aug;23(8):1490-9


Suzuki S, Nagel S, Schneider B, Chen S, Kaufmann M, Uozumi K, Arima N, Drexler HG, MacLeod RA. A second NOTCH1 chromosome rearrangement: t(9;14)(q34.3;q11.2) in T-cell neoplasia. Leukemia. 2009 May;23(5):1003-6

Tonooka Y, Fujishima M. Comparison and critical evaluation of PCR-mediated methods to walk along the sequence of genomic DNA. Appl Microbiol Biotechnol. 2009 Nov;85(1):37-43


Yin CC, Lin KI, Ketterling RP, Knudson RA, Medeiros LJ, Barron LL, Huh YO, Luthra R, Keating MJ, Abruzzo LV. Chronic lymphocytic leukemia With t(2;14)(p16;q32) involves the BCL11A and IgH genes and is associated with atypical morphologic features and unmutated IgVH genes. Am J Clin Pathol. 2009 May;131(5):663-70


This article should be referenced as such: