ELSEVIER

Cancer Genetics and Cytogenetics 159 (2005) 53–57

Cancer Genetics and Cytogenetics

Mini-Review

# Progress in concurrent analysis of loss of heterozygosity and comparative genomic hybridization utilizing high density single nucleotide polymorphism arrays

Xiaofeng Zhou[a], Nagesh P. Rao[b], Steven W. Cole[c,d,e], Samuel C. Mok[f], Zugen Chen[g], David T. Wong[a,d,e,*]

[a]*School of Dentistry, Dental Research Institute, University of California at Los Angeles, PO Box 951668, Los Angeles, CA 90095-1668*
[b]*Department of Pathology & Laboratory Medicine, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA*
[c]*Division of Hematology-Oncology, Department of Medicine, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA*
[d]*Jonsson Comprehensive Cancer Center, University of California at Los Angeles, Los Angeles, CA*
[e]*Molecular Biology Institute, University of California at Los Angeles, Los Angeles, CA*
[f]*Laboratory of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA*
[g]*Department of Human Genetics, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA*

Received 2 September 2004; received in revised form 22 September 2004; accepted 23 September 2004

**Abstract** Genetic aberrations, such as deletions and amplifications are among the major pathogenetic mechanisms underlying many medical disorders. Analysis of chromosomal aberrations is particularly important in cancer research, where amplifications of oncogenes and deletions of tumor suppressor genes are major steps in the "multi-hit" process of tumorigenesis. Genome-wide molecular biological analyses, such as loss of heterozygosity (LOH) profiling and comparative genomic hybridization (CGH) have significantly enhanced our ability to detect chromosomal aberrations in cancer cells and assess their role in tumorigenesis. The recent introduction of high-density oligonucleotide arrays for measuring single nucleotide polymorphisms (SNP) has sparked a new wave of high-resolution genetic mapping studies, including LOH and CGH applications on various cancer types. This review highlights recent progress on concurrent LOH and CGH analyses utilizing high density SNP arrays and their application in cancer research. 

## 1. Introduction

Chromosomal aberrations are characteristic of human tumors [1,2]. The most comprehensive approach for detecting genetic alteration is to strategically sequence the entire genome of each malignant specimen and compare that to the genomic sequence from matching normal tissue. This is not yet an affordable approach. However, currently available molecular genetic technologies such as comparative genomic hybridization (CGH) and loss of heterozygosity (LOH) analyses provide feasible approaches for comprehensive screening of genomic alterations with reasonably high resolution. Both CGH and LOH approaches have their unique advantages, but they also have their own limitations which have motivated efforts to combine these 2 approaches. Combined LOH and CGH analyses have been applied to renal, ovarian, bladder, and other tumors [3–5] using several different assay platforms. In this review, we will survey recent progress on combining the power of LOH and CGH in the context of newly developed high density SNP oligonucleotide arrays to provide a precise and high resolution mapping of genetic alterations.

## 2. Strengths and limitations of LOH and CGH technologies

CGH was developed to survey gene copy-number abnormalities (amplifications and deletions) across a whole genome. With CGH, differentially labeled test/disease and reference genomic DNAs are co-hybridized to normal metaphase chromosomes, and the fluorescence ratios along the

* Corresponding author. Tel: 310-206-3048; fax: 310-794-7109.
*E-mail address:* dtww@ucla.edu (D.T. Wong).

0165-4608/05/$ – see front matter 
doi:10.1016/j.cancergencyto.2004.09.014

length of chromosomes provide a cytogenetic representation of DNA copy-number variation. However, CGH has a limited mapping resolution (~20 Mb). Other high-resolution technologies, such as fluorescence in situ hybridization (FISH), are prohibitively labor-intensive on a genomic scale. Microarray-based CGH is a second generation approach in which fluorescence ratios at arrayed DNA elements provide a locus-by-locus measure of gene copy-number variation [6,7].

Although this approach has the potential to significantly increase mapping resolution, most array CGH methods have utilized large genomic clones (for example BACs) which reduce the spatial sensitivity. In addition, large genomic clones also suffer from reduced specificity due to their inclusion of common repeats [e.g., *Alu* and long interspersed nuclear elements (LINEs)], redundant sequences [e.g., low copy repeats (LCRs also known as segmental duplications)], and segments of extensive sequence similarity (pseudo genes or paralogous genes) [8]. A new generation of microarray-based CGH utilizes cDNAs arrays representing entire set of mapped human genes to provide a mapping resolution at single gene level [9,10]. However, this method also suffers from relatively low sensitivity and specificity, especially for detection of single-copy gene deletions and low-copy number gains [8]. This is believed to be attributable to biases introduced by imperfect hybridization of labeled genomic DNA to the spotted cDNA probes on the array.

Chromosomal aberrations include segments of allelic imbalance identifiable by loss of heterozygosity (LOH) at polymorphic loci. LOH is detectable by genotyping and can be used to implicate regions harboring tumor suppressor genes. Allelic losses, which are caused by mitotic recombination, gene conversion, or non-disjunction cannot be detected by CGH and thus require LOH analysis for their identification. Global LOH patterns can be generated through allelotyping of malignancies and corresponding normal tissues using polymorphic microsatellite markers or restriction fragment length polymorphisms (RFLPs). However, this approach is tedious, labor intensive, and requires large amount of sample DNA, allowing only a modest number of makers to be screened. High-density whole genome allelotyping cannot be readily performed. Furthermore, while LOH profiling can infer genomic loss, it cannot detect any amplification that might be involved in pathogenesis.

Aside from methodological drawbacks, there are also conceptual problems associated with both LOH and CGH technologies. Although CGH can detect net gain or loss of genetic materials, it does not identify situations in which the loss of one allele is followed by reduplication of the remaining allele. These latter changes would still be identifiable by LOH studies, which thus complement CGH well for this reason. On the other hand, in analyzing complicated genomes such as genomically unstable tumors, LOH results need to be interpreted cautiously since apparent LOH may be caused by events other than the loss of one allele, such as the differential amplification of the other allele. The advantages and limitations of LOH and CGH point to the need to combine these two approaches. Ideally, the most reliable survey of genomic integrity should provide both locus-specific genotypes and accurately quantify the copy number of each allele.

## 3. Concurrent LOH and CNA analysis utilizing high-density SNP arrays

SNPs are the most frequent form of DNA variation present in the human genome, and over two million SNPs have been identified (http://www.ncbi.nlm.nih.gov/SNP/). Because of their abundance, even spacing, and stability across the genome, SNPs have significant advantages over RFLPs and microsatellite markers as a basis for high-resolution whole genome allelotyping with accurate copy number measurements. SNP scoring is easily automated and high-density oligonucleotide arrays have recently been generated to support large-scale high throughput SNP analysis [11]. High-density SNP allele arrays have improved significantly and it is now possible to genotype over 10,000 SNPs using a single primer with the Affymetrix 10K SNP mapping array with a mapping resolution of approximately 210 Kb [12]. This platform was initially designed for case-control, family-based association studies, but cancer genetics researchers have quickly adapted it for LOH analysis.

Early attempts to apply SNP arrays to LOH profiling proved to be highly fruitful [13–15], yet those studies still suffered from the same conceptual limitations applying to previous LOH studies using microsatellite markers or RFLPs. However, an important recent advance has been made to utilize the hybridization intensity on the SNP arrays to generate copy number data (equivalent to the data generated from CGH). Two novel statistical bioinformatics packages (dChipSNP [16] and Affymetrix gene chip chromosome copy number tool) have been developed to simultaneously extract both genotype and hybridization intensity data for each SNP probe set. Thus, the SNP array approach offers a unique opportunity to analyze copy number abnormalities and LOH simultaneously using a single platform that provides cross-validation and complementation. Using a panel of cell lines with known genomic alterations, our recent studies have shown that this SNP array-based approach can provide highly concordant LOH and CGH analyses for the detection of discrete chromosomal amplifications or deletions [17].

In one of the cell lines studied (GM03047, Coriell Cell Repositories/NIGMS http://locus.umdnj.edu/nigms/), our analysis sub-localized the breakpoints of a deletion that was previously identified as just a 10p deletion. We identified the monosomy region to be between bands 10p12 and 10p14 (Fig. 1A). This interstitial deletion was further confirmed by fluorescence in situ hybridization (FISH) with the chromosome 10 specific sub-telomere probes (Fig. 1B). Similar

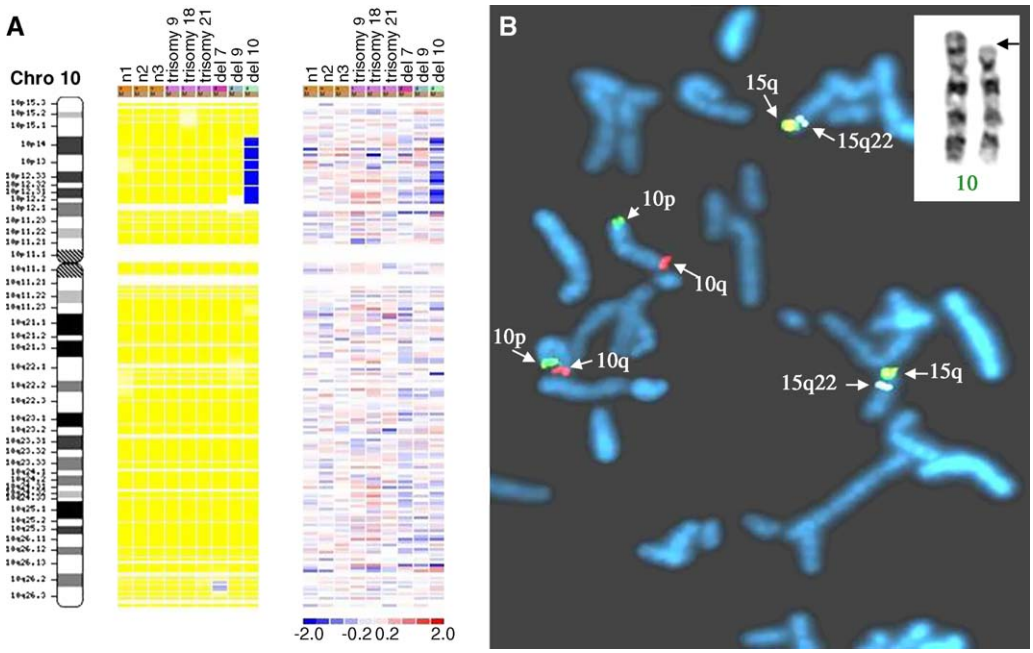


Fig. 1. Concurrent analysis of LOH and CNA on trisomy and deletion cell lines using high density SNP array. *(A)* The LOH regions and the CNA regions were detected and demarcated as described previously [13,16–18]. Each column represents one cell line, and each row represents a SNP marker. Color code for LOH profiling (right panel): blue, LOH; yellow, retained; gray, uninformative; white, no call. Copy numbers (left panel) were represented by log 2 intensity ratio (disease/reference) as indicated in the figure. The cell lines used were: normal cells (n1, n2, and n3); trisomic cells (trisomy 9, trisomy 18, and trisomy 21, with an extra copy of 9pter→q13, 18, and 21, respectively); and deletion cells (del(7), del(9), and del(10), with deletions reported previously at 7pter→q34, 9pter→p21, and 10qter→p11, respectively). *(B)* Sub-telomere FISH was performed as described previously [24] to verify the results from concurrent LOH and CGH analysis on del(10) cells by confirming that the 10p deletion was interstitial with the intact sub-telomere regions. Probes used were: 10ptel006 (10pter probe, green); 10qtel24 (10qter probe, red); PML (15q22 probe, aqua); and AFMA224XHI (15qter probe, yellow). Two normal signals for both 10p and 10q sub-telomeres were clearly identified. Insert: G-banded partial karyotype showing the del(10p).

high-density SNP array-based concurrent LOH and CGH has been successfully utilized for mapping genomic alterations in several cancer types, including breast and lung cancer [18], head and neck cancer [17], and osteosarcoma [19].

One significant advantage of concurrent LOH and CGH analysis involves their complementary abilities to more precisely define the nature of genomic alternations. For example, recent studies by our groups and by other laboratories have shown that by combining LOH and CGH analyses, it is possible to distinguish different genetic mechanisms that lead to LOH, such as hemizygous deletion (copy-reducing) or recombination/gene conversion (copy-neutral) or preferential amplification of one parental allele (copy-increasing) that underlying LOH events [17–19]. If an observed LOH is accompanied by loss of one copy (as in Fig. 1), then this LOH event is caused by hemizygous deletion. If an observed LOH is accompanied by no change of copy number, then this suggests that the LOH event is caused by copy-neutral events such as mitotic nondisjunction followed by duplication of one parental chromosomal. However, if an LOH event is accompanied by significant increase in copy numbers, then this suggests preferential amplification of one parental allele that may be masking the presence of the other allele. Alternatively, the increase in copy number in a LOH region may suggest a loss of one allele followed by amplification of the remaining allele.

Aside from the advances in high-density SNP array technology, the integration of other state-of-the-art technologies, such as laser capture microdissection and whole genome amplification of limited input DNA has further enhanced the power of concurrent LOH and CGH analysis. The integration of these technologies permits genome-wide profiling of a selected pure disease cell population from critical clinical samples, which would otherwise be impossible due to the limited cell numbers and heterogeneity of the tumor specimen. Whole genome amplification using a Phi29 polymerase-based isothermal amplification method [20] has been shown to produce a greater than 99% concordance in repeated SNP array analysis of amplified versus un-amplified original DNA samples ([21] and our unpublished results). Genomic amplification of focally captured cell samples will significantly expand the spectrum of pathological samples eligible for analysis by SNP-based LOH and CGH.

## 4. Looking ahead

Many oncogenes and tumor suppressor genes have been identified by pinpointing recurrently deleted or amplified

regions in a variety of tumors. The recent progress in concurrent LOH and CGH analysis will accelerate this ongoing process. However, the development of concurrent genomic analysis outlined in this review marks only the beginning of a new generation of high-resolution and high-throughput tools for genomic structural analysis. There is room for extensive improvements in existing approaches. The new Affymetrix 100K SNP mapping array is now available and higher density SNP array in on the way. These higher density platforms will enable a significant improvement in resolution and coverage at a reasonable cost. In addition, several groups have sought to combine DNA-based structural analyses with RNA-based expression analyses to identify the functional consequences of genomic alterations. Some recent work suggests that it might actually be possible to map genetic alterations purely from expression data using novel bioinformatics strategies [22,23]. The combination of genomic expression and structure data provides unprecedented opportunities for defining the nature of genetic alterations in disease. If the same gene shows DNA copy number abnormalities and altered expression, not only do we have independent validation of the profiling data but we will also have determined the mechanism of altered expression of that gene. Importantly, this approach provides a method for identifying causal gene expression alterations and separating them from a potentially much larger set of expression consequences that follow downstream. This will provide a significant advance in prioritizing the results of gene expression analysis for further investigation and clinical intervention. With the ongoing rapid progress in genomic mapping, transcriptional microarrays, and proteomic profiling, we envision that in the near future, concurrent analysis will be carried out at genomic (DNA), transcriptional (RNA), expressional (protein) levels as well as the level of phenotype/clinical outcome. This will have a significant impact in the field of cancer research, and may also lead to new clinical approaches for cancer and other genetic diseases diagnosis and treatments.

## Acknowledgments

This work was supported in part by NIH PHS grants R01 DE015970-01 (to D. Wong), R21 AI49135 and R01 AI52737 (to S. Cole), R33CA103595 and P50CA165009 (to S. Mok), K22 DE014847 and a TRDRP grant 13KT-0028 (to X. Zhou). The Affymetrix 10K SNP mapping array hybridization and scanning were done in the UCLA DNA microarray facility.

## References

[1] Schwab M. Oncogene amplification in solid tumors. Semin Cancer Biol 1999;9:319–25.

[2] Popescu NC, Zimonjic DB. Molecular cytogenetic characterization of cancer cell alterations. Cancer Genet Cytogenet 1997;93:10–21.

[3] Alimov A, Kost-Alimova M, Liu J, Li C, Bergerheim U, Imreh S, Klein G, Zabarovsky ER. Combined LOH/CGH analysis proves the existence of interstitial 3p deletions in renal cell carcinoma. Oncogene 2000;19:1392–9.

[4] Iwabuchi H, Sakamoto M, Sakunaga H, Ma YY, Carcangiu ML, Pinkel D, Yang-Feng TL, Gray JW. Genetic analysis of benign, low-grade, and high-grade ovarian tumors. Cancer Res 1995;55:6172–80.

[5] Obermann EC, Junker K, Stoehr R, Dietmaier W, Zaak D, Schubert J, Hofstaedter F, Knuechel R, Hartmann A. Frequent genetic alterations in flat urothelial hyperplasias and concomitant papillary bladder cancer as detected by CGH, LOH, and FISH analyses. J Pathol 2003;199:50–7.

[6] Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 1998;20:207–11.

[7] Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, Ling V, MacAulay C, Lam WL. A tiling resolution DNA microarray with complete coverage of the human genome. Nat Genet 2004;36:299–303.

[8] Mantripragada KK, Buckley PG, de Stahl TD, Dumanski JP. Genomic microarrays in the spotlight. Trends Genet 2004;20:87–94.

[9] Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 1999;23:41–6.

[10] Zhou X, Jordan RCK, Mok S, Birrer MJ, Wong DT. DNA copy number abnormality of oral squamous cell carcinoma detected by cDNA array-based CGH. Cancer Genet Cytogen 2004;151:90–2.

[11] Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 1998;280:1077–82.

[12] Matsuzaki H, Loi H, Dong S, Tsai YY, Fang J, Law J, Di X, Liu WM, Yang G, Liu G, Huang J, Kennedy GC, Ryder TB, Marcus GA, Walsh PS, Shriver MD, Puck JM, Jones KW, Mei R. Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. Genome Res 2004;14:414–25.

[13] Zhou X, Li C, Mok SC, Chen Z, Wong DT. Whole genome loss of heterozygosity profiling on oral squamous cell carcinoma by high-density single nucleotide polymorphic allele (SNP) array. Cancer Genet Cytogenet 2004;151:82–4.

[14] Lindblad-Toh K, Tanenbaum DM, Daly MJ, Winchester E, Lui WO, Villapakkam A, Stanton SE, Larsson C, Hudson TJ, Johnson BE, Lander ES, Meyerson M. Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. Nat Biotechnol 2000;18:1001–5.

[15] Mei R, Galipeau PC, Prass C, Berno A, Ghandour G, Patil N, Wolff RK, Chee MS, Reid BJ, Lockhart DJ. Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. Genome Res 2000;10:1126–37.

[16] Lin M, Wei LJ, Sellers WR, Lieberfarb M, Wong WH, Li C. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. Bioinformatics 2004;20:1233–40.

[17] Zhou X, Mok SC, Chen Z, Li Y, Wong DT. Concurrent analysis of loss of heterozygosity (LOH) and copy number abnormality (CNA) for oral premalignancy progression using Affymetrix 10K SNP mapping array. Hum Genet 2004;115:327–30.

[18] Zhao X, Li C, Paez JG, Chin K, Janne PA, Chen TH, Girard L, Minna J, Christiani D, Leo C, Gray JW, Sellers WR, Meyerson M. An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res 2004;64:3060–71.

[19] Wong KK, Tsang YT, Shen J, Cheng RS, Chang YM, Man TK, Lau CC. Allelic imbalance analysis by high-density single-nucleotide polymorphic allele (SNP) array with whole genome amplified DNA. Nucleic Acids Res 2004;32:69.

[20] Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore SF, Egholm M, Lasken RS. Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci U S A 2002;99:5261–6.

[21] Paez JG, Lin M, Beroukhim R, Lee JC, Zhao X, Richter DJ, Gabriel S, Herman P, Sasaki H, Altshuler D, Li C, Meyerson M, Sellers WR. Genome coverage and sequence fidelity of phi29 polymerase-based multiple strand displacement whole genome amplification. Nucleic Acids Res 2004;32:71.

[22] Crawley JJ, Furge KA. Identification of frequent cytogenetic aberrations in hepatocellular carcinoma using gene-expression microarray data. Genome Biol 2002;3:Research0075.

[23] Zhou X, Cole SW, Hu S, Wong DT. Detection of DNA copy number abnormality by microarray expression analysis. Hum Genet 2004; 114:464–7.

[24] Pettenati MJ, Jackle B, Bobby P, Stewart W, Von Kap-Herr C, Mowrey P, Rao PN, May KM. Unexpected retention and concomitant loss of subtelomeric regions in balanced chromosome anomalies by FISH. Am J Med Genet 2002;111:48–53.