

Allelic imbalance analysis by high-density single-nucleotide polymorphic allele (SNP) array with whole genome amplified DNA

Kwong-Kwok Wong*, Yvonne T. M. Tsang, Jianhe Shen, Rita S. Cheng, Yi-Mieng Chang, Tsz-Kwong Man and Ching C. Lau*

Texas Children's Cancer Center, Cancer Genomics Group, MC3-3320, Department of Pediatrics, Baylor College of Medicine, 6621 Fannin Street, Houston, TX 77030, USA

Received March 19, 2004; Revised and Accepted April 22, 2004

ABSTRACT

Besides their use in mRNA expression profiling, oligonucleotide microarrays have also been applied to single-nucleotide polymorphism (SNP) and loss of heterozygosity (LOH) or allelic imbalance studies. In this report, we evaluate the reliability of using whole genome amplified DNA for analysis with an oligonucleotide microarray containing 11 560 SNPs to detect allelic imbalance and chromosomal copy number abnormalities. Whole genome SNP analyses were performed with DNA extracted from osteosarcoma tissues and patient-matched blood. SNP calls were then generated by Affymetrix® GeneChip® DNA Analysis Software. In two osteosarcoma cases, using unamplified DNA, we identified 793 and 1070 SNP loci with allelic imbalance, respectively. In a parallel experiment with amplified DNA, 78% and 83% of these SNP loci with allelic imbalance was detected. The average false-positive rate is 13.8%. Furthermore, using the Affymetrix® GeneChip® Chromosome Copy Number Tool to analyze the SNP array data, we were able to detect identical chromosomal regions with gain or loss in both amplified and unamplified DNA at cytoband resolution.

INTRODUCTION

Malignant tissues frequently exhibit chromosomal aberrations and altered gene expression. The altered transcript levels in cancer genomes are often related to gene copy number changes such as amplification of oncogenes and loss of tumor suppressor genes as detected by homozygous deletion or loss of heterozygosity (LOH). In the past LOH patterns have been detected by allelotyping using restriction fragment length polymorphism (RFLP), and later by microsatellite markers (1,2). However, owing to the relative low abundance of microsatellite markers, the resolution for whole genome scanning is limited to 5 cM with commercially available sets

of primers, and the process for whole genome analysis is long and tedious. Additionally, microgram quantities of genomic DNA are needed for whole genome allelotyping. With the discovery of more than 1.4 million single-nucleotide polymorphisms (SNPs) distributed throughout the human genome at an average density of one SNP per kilobase of DNA sequence, high-resolution genome-wide allelotyping became a reality. Initially, Affymetrix® HuSNP GeneChip® containing 1494 SNPs was used to detect sequence polymorphism (1,2), and several recent studies have subsequently utilized the Affymetrix® HuSNP GeneChip® to identify regions exhibiting recurrent LOH in tumor tissues from breast, bladder, prostate and small-cell lung cancers (3–7).

Changes in DNA copy numbers are one of the characteristics of genomic instability common to many human cancers. Although comparative genomic hybridization (CGH) is an effective genome-wide technique to detect net gain or loss of genetic materials, it fails to recognize situations where there might be loss of one allele followed by reduplication of the remaining allele. These latter changes would still be identifiable by LOH studies, which complement CGH studies very well for this reason. However, in analyzing complicated genomes such as those of osteosarcomas that reflect a high level of genomic instability, LOH results need to be interpreted cautiously since some LOH at some loci may be caused by events other than the loss of one allele, such as the differential amplification of one allele. Thus the use of the term allelic imbalance to describe LOH results may be technically more accurate. Ideally, the most reliable method to characterize allelic imbalances should have the ability to not only provide locus-specific genotypes but also to quantify accurately the copy number of each allele. With more than 1.4 million SNPs already validated, high-density SNP array is a potential platform for high-resolution whole genome allelotyping with accurate copy number measurements. The high-density SNP allele array has improved significantly recently and parallel genotyping of over 10 000 SNPs using a one-primer assay is now feasible (8). The Affymetrix 10K SNP array (8) contains 11 560 SNP alleles with high frequencies of heterozygosity (average 36% based on Affymetrix in-house data). This new SNP array platform is shown to have high

*To whom correspondence should be addressed. Tel: +1 832 824 4373; Fax: +1 832 825 4038; Email: kkwong@bcm.tmc.edu
Correspondence may also be addressed to Ching C. Lau. Tel: +1 832 824 4543; Fax: +1 832 825 4038; Email: cclau@txccc.org

Table 1. Comparison of LOH calls with unamplified versus amplified total genomic DNA from two osteosarcoma patients

SNP calls	B versus T	BA versus TA	Intersection ^a	% False positives ^b	% LOH missed ^b
OST197 blood versus tumor					
AB to AA	533	522	450	13.8	15.6
AB to BB	537	500	434	13.2	19.2
OST449 blood versus tumor					
AB to AA	408	374	317	15.2	22.3
AB to BB	385	343	298	13.1	22.6

B, unamplified blood DNA; T, unamplified tumor DNA; BA, amplified blood DNA; TA, amplified tumor DNA.

^aSNP detected as LOH by both amplified (BA versus TA) and unamplified (B versus T) DNA samples.

^bThe percentage of false positive is based on the assumption that unamplified (B versus T) DNA detected LOH correctly:

% false positive = (BA versus TA – intersection)/(BA versus TA)

% LOH missed = (B versus T – intersection)/(B versus T).

accuracy (99.5%), reproducibility (99.9%) and call rate (95%) (8). The accuracy measurement is based on the concordance between SNP calls by SNP arrays and genotypes generated by the high-throughput single-base extension method or direct sequencing. Here we report the use of the 10K SNP array to perform genome-wide allelotyping with osteosarcoma tissues and validation of some of the results by parallel allelotyping studies with microsatellite markers and CGH studies.

Osteosarcoma is the most common malignant bone tumor in children and young adults (9) and is characterized by extremely complex karyotypes. Because the amount of osteosarcoma tissue that can be obtained from initial biopsies for research is very limited, we evaluated the feasibility of using whole genome amplified genomic DNA from patient samples for SNP array analysis. A Phi29 polymerase-based isothermal amplification method (10) was used to generate whole genome amplified DNA. In this study, we compared the results of using amplified DNA with those of using unamplified DNA in the detection of LOH and chromosome copy number changes in two osteosarcoma cases.

MATERIALS AND METHODS

Whole genome amplification

Fresh tissues from an initial biopsy of osteosarcoma were obtained with informed consent and snap frozen in liquid nitrogen. DNA from osteosarcoma tissue was recovered from the organic phase following TRIzol (Invitrogen, Carlsbad, CA) extraction of RNA and was further purified using DNeasy tissue Kit (Qiagen Inc., Valencia, CA). Matching normal DNA from the same patient was extracted from blood using the Wizard DNA Extraction Kit (Promega, Madison, WI). Whole genome amplification was performed with the GenomiPhi DNA amplification kit (Amersham Biosciences, Piscataway, NJ). The method employs the unique strand displacement property of Phi29 DNA polymerase (10) to amplify linear DNA. For each reaction, 9 μ l of reaction buffer and 1 μ l of Phi29 enzyme were added to 1 μ l of DNA sample containing 10 ng of genomic DNA. The reactions were allowed to proceed at 30°C for 16 h.

SNP GeneChip assay

DNA labeling, hybridization, washing and staining of the 10K SNP mapping arrays were performed according to the standard Single Primer GeneChip Mapping Assay protocol (Affymetrix Inc., Santa Clara, CA). First, 250 ng of either amplified DNA or unamplified DNA was digested with *Xba*I and then ligated to *Xba*I adaptor before subsequent PCR amplification using AmpliTaq Gold (Applied Biosystems, Foster City, CA). To obtain enough PCR products, four 100 μ l PCRs were set up for each *Xba*I adaptor-ligated DNA sample. The PCR products from the four reactions were then pooled and purified. A final 20 μ g of PCR product was fragmented with DNase I and visualized on a 4% TBE agarose gel to confirm that the sizes ranged from 50 to 100 bp. Fragmented PCR products were then end-labeled with biotin and hybridized to the array. Detection was performed with an Affymetrix Fluidics Station 400 and an Agilent GeneArray Scanner.

Microsatellite analysis

Microsatellite analyses were performed with primers and reagents from the ABI PRISM[®] Linkage Mapping Set Version 2.5. In a 15 μ l reaction, 15 ng of DNA was combined with 9 μ l of true allele mix and 1 μ l of primer pairs. The following PCR conditions were used: 1 cycle of 95°C for 12 min, then 10 cycles of 94°C for 15 s, 55°C for 15 s and 72°C for 30 s, followed by 20 cycles of 89°C for 15 s, 55°C for 15 s and 72°C for 30 s. Finally the reaction was held at 72°C for 40 min before electrophoresis on an ABI Prism 377 Sequencer. The bands were scanned and data were analyzed by ABI PRISM[®] Genescan/Genotyper software.

Data analysis

The signal intensity data from the GeneChip Operating software were analyzed by GeneChip DNA Analysis Software (GDAS). The GDAS Mapping Algorithm uses a model-based approach (11,12) to perform allele calling for all SNPs on GeneChip[®] 10K Mapping Arrays. Information about the linear chromosome location, upstream and downstream associated microsatellite markers and genes for each SNP was extracted directly from NetAffx[™] Analysis Center (<http://www.affymetrix.com>) (13). Data from eight array

Table 2. Concordance between microsatellite markers with LOH and associated SNPs

Microsatellite marker with LOH	Concordance	Associated SNP ^a	SNP associated genes ^b
D10S548	Yes	SNP_A-1507719 (354352)	FLJ30499 (0)
	Yes	SNP_A-1514467 (275041)	FLJ30499 (138820)
D10S547	Yes	SNP_A-1515335 (73967)	CUGBP2 (435332)
	Yes	SNP_A-1512522 (28578)	CUGBP2 (538238)
	Yes	SNP_A-1512797 (402177)	CUGBP2 (911837)
	Yes	SNP_A-1509048 (28484)	CUGBP2 (538144)
	Yes	SNP_A-1507528 (71842)	CUGBP2 (337457)
D10S591	Yes	SNP_A-1515840 (50139)	LoopADR (57865)
	Yes	SNP_A-1507999 (802983)	PITRM1 (305939); COPEB (291241)
D11S935	Yes	SNP_A-1517705 (96736)	EST (0)
D1S252	Yes	SNP_A-1516337 (795188)	FLJ38716 (0)
	Yes	SNP_A-1516514 (389864)	CD58 (0)
D1S2842	Yes	SNP_A-1510676 (66659)	FLJ40773 (117902)
D6S287	Yes	SNP_A-1514474 (733220)	BP75 (167748)
	Yes	SNP_A-1513072 (807375)	MAN1A1 (590044)
D6S446	Yes	SNP_A-1514409 (35997)	DLL1 (2647)
D8S1771	Yes	SNP_A-1513125 (400216)	EBF2 (0)
D8S258	Yes	SNP_A-1518093 (249174)	LZTS1 (5190)
D9S286	Yes	SNP_A-1514444 (214489)	PTPRD (49070)
	Yes	SNP_A-1513014 (602303)	PTPRD (0)
D1S2797	Yes	SNP_A-1516178 (939472)	TESK2 (0)
D1S2878	Yes	SNP_A-1510747 (27247)	EST (0)
	Yes	SNP_A-1517277 (27387)	EST (0)
D1S468	Yes	SNP_A-1518557 (517234)	MOT8 (399588)
D3S1271	No	SNP_A-1519022 (61676)	FLJ21551 (84348)
D4S406	No	SNP_A-1517547 (3265)	LCE (150071); ENPEP (127645)
D6S289	No	SNP_A-1512627 (103649)	JMJ (168238)
D7S798	No	SNP_A-1511326 (21696)	ARP3BETA (137620)

^aThe number in parentheses is the distance (base pairs) between the microsatellite marker and the corresponding SNP.

^bThe number in parentheses is the distance (base pairs) between the adjacent gene and the SNP.

experiments were collated using OmniViz software (<http://omniviz.com>), and all SNPs with LOH (genotype changing from AB in the normal blood DNA to AA or BB in the corresponding tumor DNA) were identified using a dynamic query tool within the OmniViz software package. Individual SNP copy numbers and chromosomal regions with gains or losses were evaluated with the Affymetrix® GeneChip® Chromosome Copy Number Tool.

RESULTS AND DISCUSSION

Whole genome amplification of osteosarcoma DNA

Since a very limited amount of tissue is available for research from initial biopsies of osteosarcoma, we evaluated the feasibility of using whole genome amplified DNA to study changes in DNA copy number in osteosarcoma by using a high-density SNP array. Tumor DNA recovered after TRIzol extraction of RNA was used for whole genome DNA amplification using a Phi29 polymerase-based GenomiPhi Kit, making it theoretically possible to perform both expression and DNA analyses with the same piece of tissue when tissue quantity is limited. However, we found that a DNeasy cleaning step is necessary in order to obtain robust amplification with TRIzol-extracted DNA. Typical yield of amplified DNA from the GenomiPhi Kit is about 3–5 µg from 10 ng of genomic DNA as starting material. We amplified DNA from two osteosarcoma samples and the corresponding normal DNA from blood. Amplified DNA was evaluated for

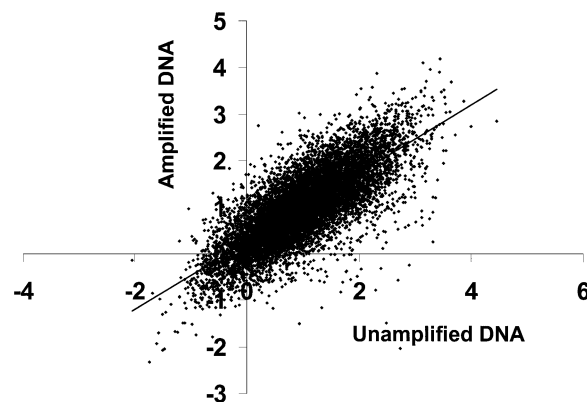


Figure 1. Correlation between the estimated copy number of 11 560 SNPs from amplified and unamplified osteosarcoma DNA. The copy number is estimated by the Affymetrix® GeneChip® Chromosome Copy Number Tool and is in log₂ scale. The Pearson correlation coefficient is 0.77.

allelotyping using both microsatellite markers and SNP arrays. Initially, a total of 78 pairs of primers, representing 78 microsatellite markers, were selected from the ABI PRISM® Linkage Mapping Set Version 2.5. These microsatellite markers were used to compare a pair of amplified and unamplified control DNA. Both amplified and unamplified DNA gave the same allelic calls for 76 out of 78 loci tested. Next, amplified normal and tumor DNAs from osteosarcoma patients were used for microsatellite allelotyping using the 76

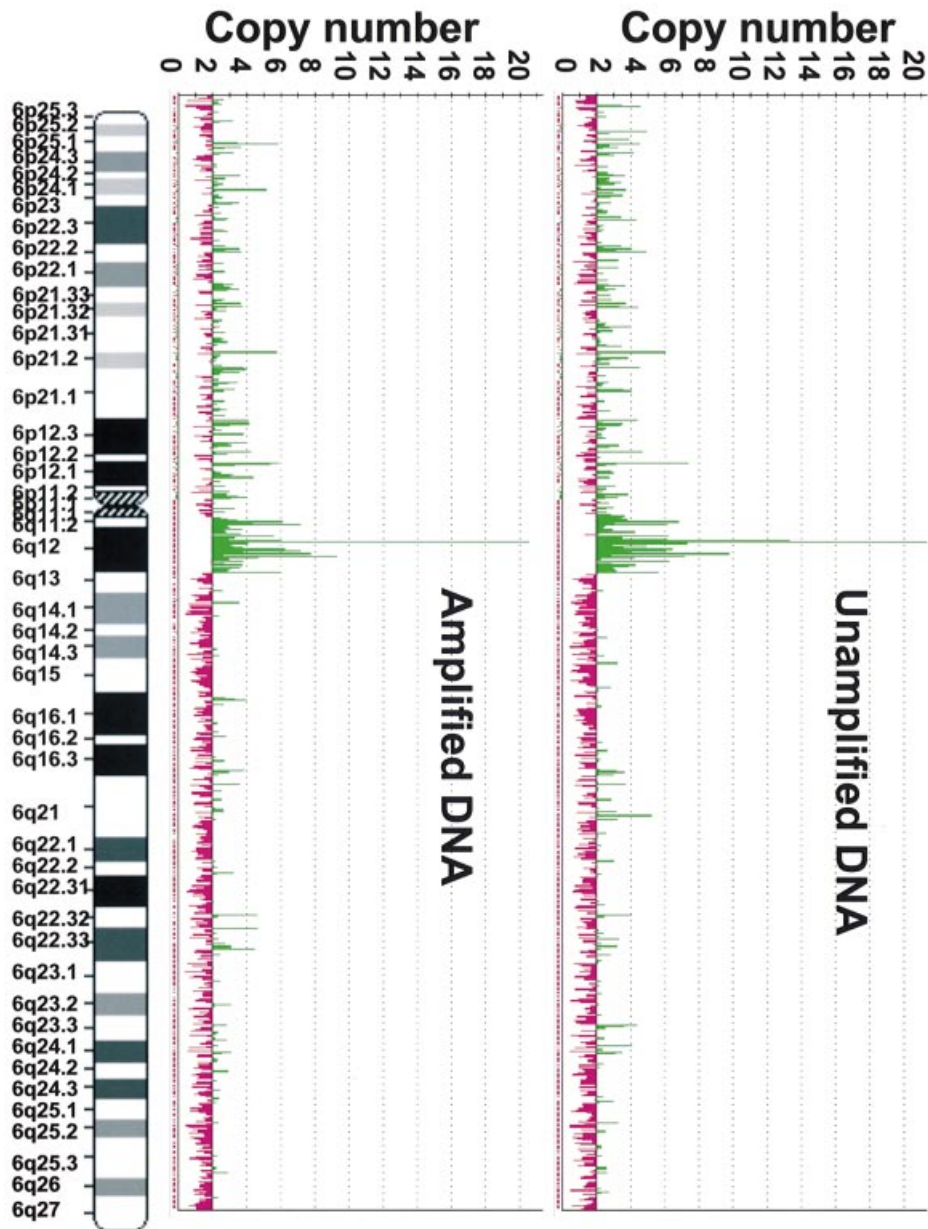


Figure 2. Copy number of individual SNPs in chromosome 6 detected with amplified and unamplified tumor DNA from case OST197. The copy number results are plotted using two colors: green for values above the threshold (2) and magenta for values below the threshold. Included in the graph is a representation of the genotype calls associated with the SNPs (small color bars to the right of the ideogram). Green represents heterozygous calls while magenta represents homozygous calls. An ideogram showing the corresponding cytoband locations of each SNPs on chromosome 6 is aligned in the bottom of the graph.

loci validated for whole genome amplification. In one case (OST197), 46 out of 76 microsatellite markers were found to have LOH.

Accuracy of LOH detection by SNP array with amplified DNA

To evaluate the accuracy of LOH detection using amplified DNA, a total of eight SNP arrays were used to analyze two independent osteosarcoma cases (OST197 and OST449) with corresponding patient-matched blood samples. We were able to obtain genotype calls (AA, BB or AB) from 86 to 97% (average $92.13 \pm 4.43\%$) of the 11560 SNPs in the SNP array using both amplified DNA and unamplified DNA derived from

the four independent samples. Moreover, the difference between the mean call rates of using unamplified and amplified DNA is insignificant based on a *t* test on the four paired samples ($p = 0.228$). The SNP call for each of the allelic loci was determined by the GeneChip® DNA Analysis Software. More than 3900 SNP loci had heterozygous calls (AB) in both the amplified and unamplified blood DNA, which are informative for determining LOH in the corresponding tumor DNA. Table 1 summarizes the LOH results using both amplified and unamplified DNA from two cases of osteosarcoma.

A total of 1070 and 793 SNP loci with LOH were detected in OST197 and OST449, respectively, with unamplified DNA.

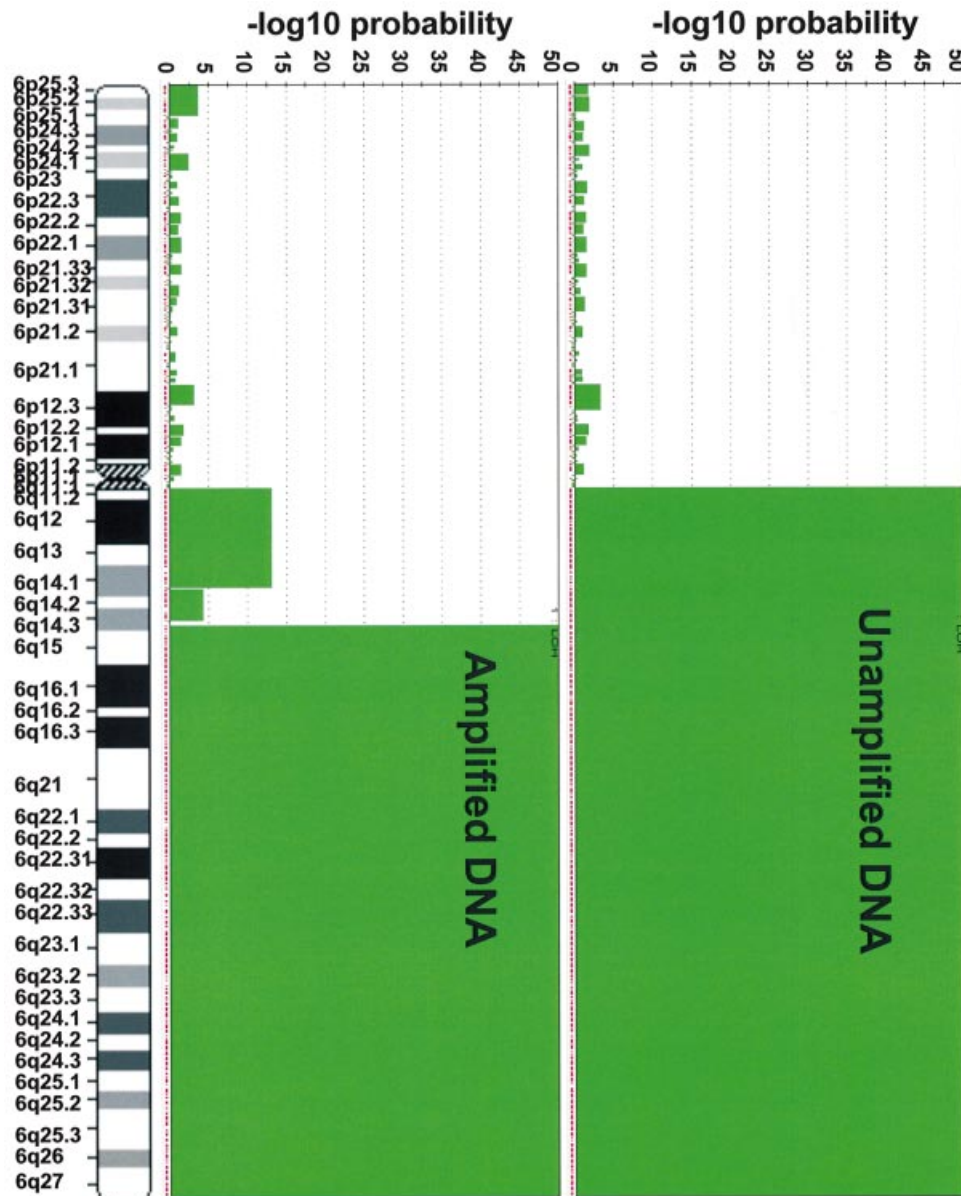


Figure 3. Probability plot of LOH calls in amplified and unamplified tumor DNA without reference to the normal DNA. An example for chromosome 6 of OST197 is shown.

On the other hand, 1022 and 717 SNP loci with LOH were detected in OST197 and OST449, respectively, with amplified DNA. More than 78% of the LOH detected with unamplified DNA were also detected using amplified DNA. The average false-positive and false-negative rates for using amplified DNA were 13.5% and 19.9% respectively. The average accuracy is calculated as 81% (CI, 79% and 82%) at 95% significance (based on exact method and binomial confidence interval) by SPSS (SPSS Inc., Chicago, IL).

Concordance between LOH detection by microsatellite markers and SNP array

Since we have identified 46 chromosomal loci with LOH in OST197 by allelotyping with microsatellite markers, we evaluated the concordance between LOH detection by

microsatellite markers and SNP arrays in this case. Using the NetAffx™ Analysis Center, we were able to identify SNPs that are within 1 Mb for 31 of these 46 microsatellite markers, but only 18 microsatellite markers have associated informative SNPs. Since the average heterozygosity of microsatellite marker (0.79) from the ABI PRISM® Linkage Mapping Set Version 2.5 is much higher than that of the SNPs in the array (0.36), it is conceivable that some of these microsatellite markers do not have associated informative SNPs for comparison. Table 2 lists the microsatellite markers loci with LOH and the associated informative SNPs. Among these 18 microsatellite markers, 14 have associated SNPs with LOH. However, SNPs associated with the remaining four microsatellite markers do not show any LOH. Three of these microsatellite markers (D3S1271, D6S289 and D7S798) are

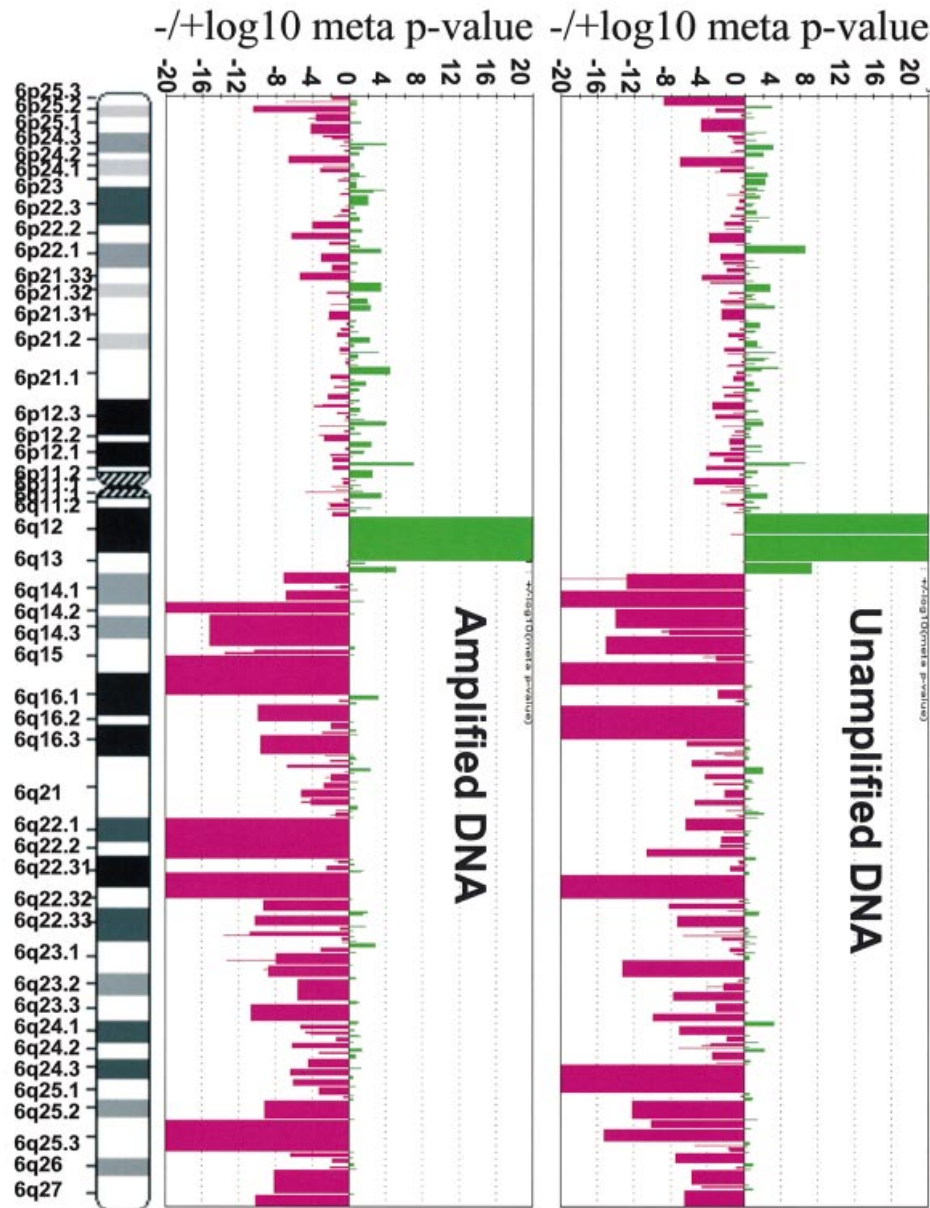


Figure 4. Significance graph of detecting copy number changes in amplified and unamplified genomic DNA. An example for chromosome 6 of OST197 is shown.

20–100 kb away from the associated SNPs and one (D4S406) is within 4 kb. There are two possible explanations for the discrepancies involving these four microsatellite markers: (i) the boundary of LOH is located between the SNP locus and the corresponding microsatellite marker, and (ii) errors of the SNP array in making heterozygous calls in tumor DNA. Sequencing these loci will help clarify the basis of the discrepancies.

SNP array is able to detect chromosomal regions with LOH and copy number changes in amplified tumor DNA without reference to normal DNA

Since patient-matched normal DNA is not always available as a reference for high-resolution allelotyping, especially in retrospective studies, we tested the feasibility of detecting

LOH or copy number changes using tumor DNA only. The Affymetrix® GeneChip® Chromosome Copy Number Tool estimates the copy number of individual SNPs by comparing the signal intensity of each SNP from the tumor sample with the mean of the corresponding SNP in a reference set containing >100 normal individuals. The estimated copy number of 11560 SNPs from unamplified OST449 DNA was plotted against that of amplified OST449 DNA (Fig. 1). The Pearson correlation of estimated SNP copy numbers between unamplified and amplified DNA is 0.77 for OST449 (Fig. 1) and 0.70 for OST197. Figure 2 is a detailed plot of the copy number for individual SNP loci along chromosome 6 from OST197. Based on the estimated copy numbers, SNP loci at 6q12–13 have 5- to 20-fold amplification (Fig. 2) as detected by using either amplified or unamplified DNA. This is

consistent with our chromosome CGH result that 6q12–13 is an amplified region (data not shown).

Figure 3 compares the detection of LOH between amplified and unamplified DNA using chromosome 6 as an illustration. The Affymetrix® GeneChip® Chromosome Copy Number Tool calculates the probability of homozygosity of each SNP using a large reference set containing >100 individuals. By treating each SNP independently, the probability that a contiguous stretch of homozygous calls happens by chance is then calculated. Chromosome regions with LOH are inferred from this probability based on the assumption that the occurrence of long stretches of homozygous regions along a chromosome is very unlikely and therefore may indicate a region of LOH. This is illustrated in Figure 3 in which $-\log_{10}$ of the probability of a given stretch of homozygous calls is plotted against the position of individual SNPs along the chromosome. It shows a comparison of LOH calls using this algorithm from both amplified and unamplified tumor DNA. Both amplified and unamplified DNA gave identical results that there is LOH of the entire 6q. With amplified DNA, the probability of LOH at 6q12–13 loci is lower than that of unamplified DNA (Fig. 3). This is because a few SNP loci at 6q12–13 were called heterozygous in amplified DNA but were called homozygous in unamplified DNA. However, these few heterozygous calls do not affect the confidence that 6q12–13 is within a LOH region. The probability that there is no LOH in this region is less than 1 in 10^{12} .

SNP loci at 6q12–13 are within a region of LOH (Fig. 3) but also have significant increase in copy number (Fig. 2). The increase in copy number in a LOH region may suggest the loss of an allele followed by amplification of the remaining allele. The ability to make such an inference is one of the advantages of SNP array over other microarray-based methods such as the use of cDNA and BAC (14,15) for allelic imbalance analysis.

Besides estimating the copy numbers for individual SNP locus, we also used the Affymetrix® GeneChip® Chromosome Copy Number Tool to evaluate the statistical significance of copy number changes along the chromosomes. A p value for each SNP signal intensity of the tumor DNA is calculated from a distribution of SNP intensities from a reference set containing >100 normal individuals. A smaller p value for a SNP locus will imply a higher significant gain or loss at that locus. The p value is \log_{10} transformed and plotted along the corresponding chromosome. Figure 4 is a comparison of the significance plot of chromosome gain and loss for chromosome 6 using both amplified and unamplified DNA. The sign of the \log_{10} -transformed p value is modified so that a positive sign will indicate chromosomal gain while a negative sign will indicate chromosomal loss. There is good concordance between amplified and unamplified DNA, such that both indicate a loss of 6q14–q27 and gain at 6q12–13 in the case of OST197. The Pearson correlation between the p value of each SNP from the amplified and unamplified DNA is 0.716. Both the gain and loss of these two regions in chromosome 6 are confirmed by CGH (data not shown).

In summary, we were able to detect allelic imbalances with whole genome amplified DNA from two osteosarcoma samples by using a high-density SNP array. Both amplified and unamplified DNA gave similar results in terms of SNP calls, LOH and chromosome copy number changes. Our data indicate that copy number changes and LOH can be estimated

or inferred from 10K SNP array data using only amplified tumor DNA with the Affymetrix® GeneChip® Chromosome Copy Number Tool. The results are comparable to that of unamplified tumor DNA (Figures 1–4). Because of the high density of the SNP array, with a median inter-SNP distance of 105 kb, localizing the associated genes near the adjacent SNPs for further analysis is made much easier (Table 2). The identification of these candidate genes will further enhance the understanding of the biology of osteosarcoma. The strategies outlined in this report can also be applied to the study of other human cancers that have limited quantities of tissue available such as those that are traditionally diagnosed by needle biopsies or tissues that require microdissection.

ACKNOWLEDGEMENTS

This work was supported in part by NIH PHS grant CA81465, and by grants from the John S. Dunn Research Foundation, the Robert J. Kleberg Jr and Helen C. Kleberg Foundation, and the Gillson Longenbaugh Foundation.

REFERENCES

1. Wang, V.W., Bell, D.A., Berkowitz, R.S. and Mok, S.C. (2001) Whole genome amplification and high-throughput allelotyping identified five distinct deletion regions on chromosomes 5 and 6 in microdissected early-stage ovarian tumors. *Cancer Res.*, **61**, 4169–4174.
2. Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X.C., Stern, D., Winkler, J., Lockhart, D.J., Morris, M.S. and Fodor, S.P. (1996) Accessing genetic information with high-density DNA arrays. *Science*, **274**, 610–614.
3. Lindblad-Toh, K., Tanenbaum, D.M., Daly, M.J., Winchester, E., Lui, W.O., Villapakkam, A., Stanton, S.E., Larsson, C., Hudson, T.J., Johnson, B.E. *et al.* (2000) Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. *Nat. Biotechnol.*, **18**, 1001–1005.
4. Schubert, E.L., Hsu, L., Cousens, L.A., Glogovac, J., Self, S., Reid, B.J., Rabinovitch, P.S. and Porter, P.L. (2002) Single nucleotide polymorphism array analysis of flow-sorted epithelial cells from frozen versus fixed tissues for whole genome analysis of allelic loss in breast cancer. *Am. J. Pathol.*, **160**, 73–79.
5. Dumur, C.I., Dechsukhum, C., Ware, J.L., Cofield, S.S., Best, A.M., Wilkinson, D.S., Garrett, C.T. and Ferreira-Gonzalez, A. (2003) Genome-wide detection of LOH in prostate cancer using human SNP microarray technology. *Genomics*, **81**, 260–269.
6. Hoque, M.O., Lee, C.C., Cairns, P., Schoenberg, M. and Sidransky, D. (2003) Genome-wide genetic characterization of bladder cancer: a comparison of high-density single-nucleotide polymorphism arrays and PCR-based microsatellite analysis. *Cancer Res.*, **63**, 2216–2222.
7. Primdahl, H., Wikman, F.P., von der Maase, H., Zhou, X.G., Wolf, H. and Orntoft, T.F. (2002) Allelic imbalances in human bladder cancer: genome-wide detection with high-density single-nucleotide polymorphism arrays. *J. Natl Cancer Inst.*, **94**, 216–223.
8. Matsuzaki, H., Loi, H., Dong, S., Tsai, Y.-Y., Fang, J., Law, J., Di, X., Liu, W.-M., Yang, G., Liu, G. *et al.* (2004) Parallel genotyping of over 10 000 SNPs using a one-primer assay on a high-density oligonucleotide array. *Genome Res.*, **14**, 414–425.
9. Lau, C.C., Harris, C.P., Lu, X.Y., Perlaky, L., Gogineni, S., Chintagumpala, M., Hicks, J., Johnson, M.E., Davino, N.A., Huvos, A.G. *et al.* (2004) Frequent amplification and rearrangement of chromosomal bands 6p12–p21 and 17p11.2 in osteosarcoma. *Genes Chromosom. Cancer*, **39**, 11–21.
10. Dean, F.B., Hosono, S., Fang, L., Wu, X., Faruqi, A.F., Bray-Ward, P., Sun, Z., Zong, Q., Du, Y., Du, J. *et al.* (2002) Comprehensive human genome amplification using multiple displacement amplification. *Proc. Natl Acad. Sci. USA*, **99**, 5261–5266.
11. Kennedy, G.C., Matsuzaki, H., Dong, S., Liu, W.M., Huang, J., Liu, G., Su, X., Cao, M., Chen, W., Zhang, J. *et al.* (2003) Large-scale genotyping of complex DNA. *Nat. Biotechnol.*, **21**, 1233–1237.

12. Liu, W.M., Di, X., Yang, G., Matsuzaki, H., Huang, J., Mei, R., Ryder, T.B., Webster, T.A., Dong, S., Liu, G. *et al.* (2003) Algorithms for large-scale genotyping microarrays. *Bioinformatics*, **19**, 2397–2403.
13. Liu, G., Loraine, A.E., Shigeta, R., Cline, M., Cheng, J., Valmeekam, V., Sun, S., Kulp, D. and Siani-Rose, M.A. (2003) NetAffx: Affymetrix probesets and annotations. *Nucleic Acids Res.*, **31**, 82–86.
14. Pollack, J.R., Perou, C.M., Alizadeh, A.A., Eisen, M.B., Pergamenschikov, A., Williams, C.F., Jeffrey, S.S., Botstein, D. and Brown, P.O. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.*, **23**, 41–46.
15. Ishkanian, A.S., Malloff, C.A., Watson, S.K., DeLeeuw, R.J., Chi, B., Coe, B.P., Snijders, A., Albertson, D.G., Pinkel, D., Marra, M.A. *et al.* (2004) A tiling resolution DNA microarray with complete coverage of the human genome. *Nat. Genet.*, **36**, 299–303.