

## 1 DESIGN OF DOP-PCR 32K BAC-BASED HUMAN GENOME ARRAY

The array was constructed using a human BAC minimal tiling clone set which was purchased from CHORI (<http://bacpac.chori.org/pHumanMinSet.htm>). This clone set provides a complete coverage of the sequenced parts of the human genome (May 2004 assembly of the Human genome (Build35/hg17)) with an average resolution of <100 kb. DNA isolated from the clones was DOP-PCR amplified (Fiegler *et al.*, 2003) and spotted on high density CodeLink slides (GE Healthcare) with a high performance printer constructed by the bioinstrumentation group at the Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA. Prior to analysis the array was validated extensively using previously well characterized samples (Diaz de Ståhl *et al.* (2007); Bruder *et al.* (2007); Piotrowski *et al.*, in preparation).

## 2 HYBRIDIZATION, SCANNING AND DATA ANALYSIS

The methods used for blocking CodeLink slides, DNA labeling, and post-hybridization processing have been described in detail elsewhere (Mantripragada *et al.*, 2006). Briefly, the slides were blocked using sodium borohydride treatment, and prehybridized (5x SSC, 0.1% SDS, 0.4% BSA; 1 h, 45°C). Test and reference DNA was labeled by random priming with Cy3 dCTP (PA 53021, GE HealthCare, UK) and Cy5 dCTP (PA55021, GE HealthCare, UK), respectively, using Bioprime Array CGH Genomic Labeling System with purification module (18095-011, Invitrogen, Carlsbad, CA). Labeled DNA was suspended in the hybridization buffer (2x SSC, 4% SDS, 50% formamide, 10% dextran sulfate) and hybridized to the array (20 h, 45°C). Post-hybridization processing consisted of four washing steps: i) 2xSSC, 0.1% SDS, 25% formamide; 20 min., 45°C; ii) 1x PBS; 10 min., RT; iii) 0.2x SSC; 15 s, RT; iv) deionized water; 5 s, RT. The slides were dried with compressed air. Image acquisition was performed using the GenePix 4000B scanner (Axon Instruments Inc, Union City, CA). Analysis of hybridization intensity was carried out using the GenePix Pro image analysis software (Axon Instruments).

Subsequent data processing steps were performed in the Linnaeus Center for Bioinformatics Data Warehouse (Ameur *et al.*, 2006). The steps of data processing were as follows: filtering out the spots containing >5% of over-saturated pixels, normalizing of data using print-tip LOESS (Yang *et al.*, 2002), discarding the spots which in both Cy3 and Cy5 channels showed signal-to-noise ratio less than three, and filtering out the measurement points which were marked *bad* or *empty* in GenePix Pro.

## REFERENCES

- Ameur, A., Yankovski, V., Enroth, S., Spjuth, O., and Komorowski, J. (2006). The LCB Data Warehouse. *Bioinformatics*, **22**(8), 1024–1026.
- Bruder, C. E., Piotrowski, A., Gijssbers, A., Andersson, R., Erickson, S., Diaz de Ståhl, T., Menzel, U., Sandgren, J., von Tell, D., Poplawski, A., Crowley, M., Crasto, C., Partridge, E. C., Tiwari, H., Allison, D. B., Komorowski, J., van Ommen, G.-J. B., Boomsma, D. I., den Dunnen, J. T., Pedersen, N. L., Wirdefeldt, K., and Dumanski, J. (2007). Phenotypically concordant and discordant monozygotic twins display different DNA copy number variation profiles. *Am J Hum Genet.* in press.
- Diaz de Ståhl, T., Sandgren, J., Piotrowski, A., Nord, H., Andersson, R., Menzel, U., Bogdan, A., Thuresson, A. C., Poplawski, A., von Tell, D., Hansson, C. M., Elshafie, A. I., Elghazali, G., Imreh, S., Nordenskjöld, M., Upadhyaya, M., Komorowski, J., Bruder, C. E., and Dumanski, J. P. (2007). Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array. *Hum Mutat.* Epub ahead of print.
- Fiegler, H., Carr, P., Douglas, E. J., Burford, D. C., Hunt, S., Scott, C. E., Smith, J., Vetrie, D., Gorman, P., Tomlinson, I. P., and Carter, N. P. (2003). DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer*, **36**(4), 361–374.
- Mantripragada, K. K., Thuresson, A. C., Piotrowski, A., Diaz de Ståhl, T., Menzel, U., Grigelionis, G., Ferner, R. E., Griffiths, S., Bolund, L., Mautner, V., Nordling, M., Legius, E., Vetrie, D., Dahl, N., Messiaen, L., Upadhyaya, M., Bruder, C. E., and Dumanski, J. P. (2006). Identification of novel deletion breakpoints bordered by segmental duplications in the NF1 locus using high resolution array-CGH. *J Med Genet*, **43**(1), 28–38.
- Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., and Speed, T. P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucl. Acids Res.*, **30**(4), e15–.