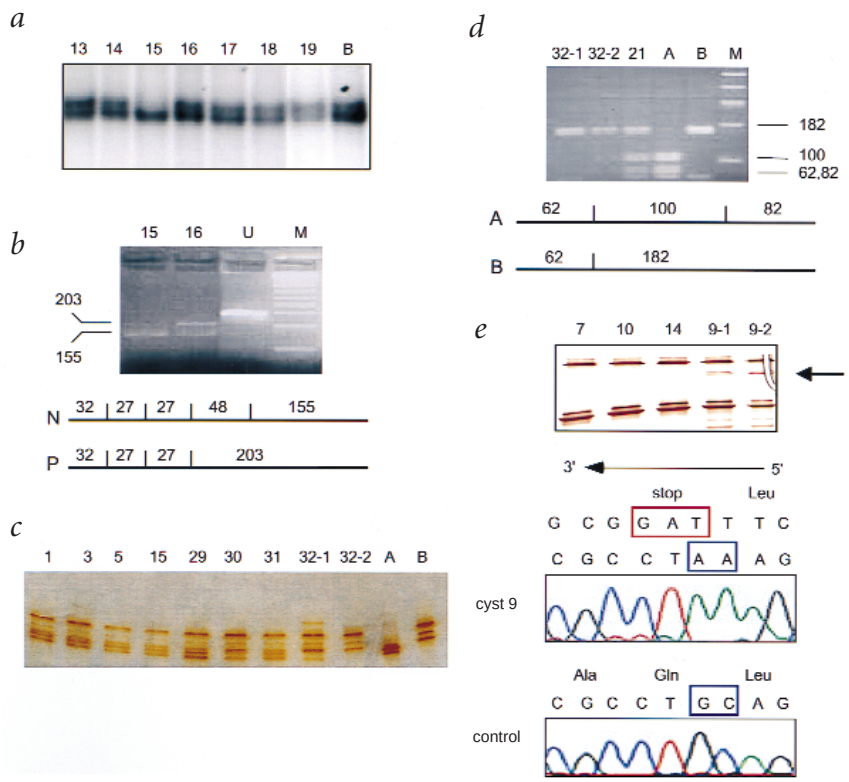

Fig. 1 Somatic mutations of *PKD1* in *PKD2* cysts. ***a***, SSCP analysis of the C5383T SNP in exon 15. A duplex pattern is seen in all samples except C15. C15 also had LOH at another SNP locus in exon 13 (T3274C; data not shown; ref. 14). ***b***, Confirmation of LOH by restriction digestion. The C5383T polymorphism (P) results in the loss of an *Hae*III site and creation of a new 203-bp fragment. C15 lacks the 203-bp band compared with a control cyst (C16). ***c***, SSCP analysis of the T7376C SNP in exon 17 of UT1270. The SSCP pattern for individuals homozygous for each allele are included (A, B; refs 2,14). All cysts have both alleles, except C32, which has only B (32-1). This result was confirmed using a second, independently amplified long-range PCR product (32-2). C15, which exhibited LOH for markers in exon 15 (*a*), is heterozygous for this marker. ***d***, Confirmation of LOH in exon 17 by restriction digestion. The T7376C substitution creates a new *Mva*II site in A (present in the control cyst, C21), which is lost in C32-1 and C32-2. U, uncut product; M, 1-kb ladder. ***e***, C9 from UT1500 has a 2-bp substitution (G1501T, C1502T; blue box) that results in a premature stop codon (red box). The exon 6 SSCP pattern for four cysts of UT1500 is shown, and the arrow identifies a variant seen in two independent PCR reactions from C9 (9-1, 9-2). The sequence tracings for normal and variant bands are shown in the 3´-5´ direction. The complementary sequence is also given with the predicted peptide sequence. Methods for *PKD1* mutation detection have been described[2,13,14].

Although the scope of our study has been limited by the restricted supply of suitable human tissues, our results suggest that somatic mutation of *PKD1* may be a modifier of disease severity in ADPKD2. We speculate that *trans*-heterozygous loss of function at the somatic level may be sufficient to disrupt the signalling pathway in which these proteins participate.

Acknowledgements

We thank the individuals with ADPKD and their families for participation. This work was supported by the NIH (DK48006, DK02562), the Polycystic Kidney Disease Research Foundation and the Kidney Foundation of Canada. G.G.G. is the Irving Blum Scholar of the Johns Hopkins University School of Medicine.

Terry Watnick[1]*, Ning He[2]*, Kairong Wang[2], Yan Liang[3], Patrick Parfrey[4], Donna Hefferton[4], Peter St George-Hyslop[3], Gregory Germino[1] & York Pei[2]

**These authors contributed equally to this work.*

[1]Department of Medicine, Division of Nephrology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [2]Department of Medicine, University Health Network, Toronto, Ontario, Canada. [3]Tanz Research Institute, Toronto, Ontario, Canada. [4]Division of Nephrology, Department of Medicine, Health Sciences Center, St Johns, Newfoundland, Canada. Correspondence should be addressed to G.G. (e-mail: ggermino@welch.jhu.edu) or Y.P. (York.Pei@uhn.on.can).

1. Qian, F., Watnick, T.J., Onuchic, L.F. & Germino, G.G. *Cell* **87**, 979–987 (1996).
2. Watnick, T.J. *et al. Mol. Cell* **2**, 247–251 (1998).
3. Koptides, M., Hadjimichael, C., Koupepidou, P., Pierides, A. & Deltas, C.C. *Hum. Mol. Genet.* **8**, 509–513 (1999).
4. Pei, Y. *et al. J. Am. Soc. Nephrol.* **10**, 1524–1529 (1999).
5. Lu, W. *et al. Nature Genet.* **17**, 179–181 (1997).
6. Wu, G. *et al. Cell* **93**, 177–188 (1998).
7. Geng, L. *et al. J. Clin. Invest.* **98**, 2674–2682 (1996).
8. Ong, A.C. *et al. Am. J. Pathol.* **154**, 1721–1729 (1999).
9. Ong, A.C. & Harris, P.C. *Lancet* **349**, 1039–1040 (1997).
10. Pei, Y. *et al. J. Am. Soc. Nephrol.* **9**, 1852–1860 (1998).
11. Torra, R. *et al. Kidney Int.* **56**, 28–33 (1999).
12. Qian, F. *et al. Nature Genet.* **16**, 179–183 (1997).
13. Watnick, T.J. *et al. Hum. Mol. Genet.* **6**, 1473–1481 (1997).
14. Watnick, T.J. *et al. Am. J. Hum. Genet.* **65**, 1561–1571 (1999).
15. Koptides, M., Mean, R., Demetriou, K., Pierides, A. & Deltas, C.C. *Hum. Mol. Genet.* **12**, 447–452 (2000).

Quantitative mapping of amplicon structure by array CGH identifies *CYP24* as a candidate oncogene

We show here that quantitative measurement of DNA copy number across amplified regions using array comparative genomic hybridization[1–4] (CGH) may facilitate oncogene identification by providing precise information on the locations of both amplicon boundaries and amplification maxima. Using this analytical capability, we resolved two regions of amplification within an approximately 2-Mb region of recurrent aberration at 20q13.2 in breast cancer. The putative oncogene *ZNF217* (ref. 5) mapped to one peak, and *CYP24* (encoding vitamin D 24 hydroxylase), whose overexpression is likely to lead to abrogation of growth control mediated by vitamin D (ref. 6), mapped to the other.

Positional localization of candidate oncogenes often includes assembly of a contig of large-insert genomic clones, such as BACs, across the region. We have shown previously that such clone sets can be used as probes to map the boundaries of chromosomal rearrangements by fluorescent *in situ* hybridization (FISH) with a finer resolution than the insert size of the clones[7,8]. Here we report that similar sub-clonal resolution mapping of varia-

tions in DNA copy number can be performed on sets of overlapping clones using array CGH. In array CGH (refs 1–4), total genomic DNA from a tumour and a normal cell population are labelled with different fluorochromes and hybridized to arrayed clones. The ratio of the fluorescence intensities on each spot in the array is then proportional to the copy number of the corresponding sequences in the tumour. As the ratio on a clone represents the average DNA copy number over its length, comparison of ratios on overlapping clones should allow aberrations to be mapped to a fraction of the length of a clone.

We demonstrated the biological value of this 'high-resolution' form of array CGH using an array comprised of overlapping BAC and P1 clones (Fig. 1*a*) to map amplicon boundaries and copy-number profiles across a 2-Mb region of recurrent amplification at 20q13.2 in breast cancer. We found that most breast tumours and cell lines with amplification at 20q13.2 (refs 2,5, and unpublished data) showed elevated and slightly varying copy number across the entire region (data not shown) clearly distinct from normal specimens (Fig. 1*b*). In three tumours, however, we observed a dramatic variation in copy number across the region (Fig. 1*c*). The profiles showed narrow regions of copy-number maxima as well as amplicon boundaries that indicate two regions that may be independently amplified in breast tumours within this 2-Mb genomic segment.

The peaks in the copy number profiles suggest that progressive 'trimming' of the amplified segment may have taken place in these tumours as copy number increased, a phenomenon that has been described in cultured cells under drug selection[9]. There is therefore an expectation that the 'driver' genes are located at these relatively narrow regions of highest copy number. The putative oncogene *ZNF217* (ref. 5) maps to the peak of the more proximal region (Fig. 1). *CYP24* is located at the more distal region, and is thus a candidate 'driver' gene for it. *CYP24* has not been previously implicated in breast cancer, but its known function supports its candidacy. *CYP24* limits the biological activity of the vitamin D signalling system[6,10,11], so that overexpression due to amplification may abrogate vitamin-D–mediated growth control. As transcription of *CYP24* is closely coupled to the level and activity of the vitamin D

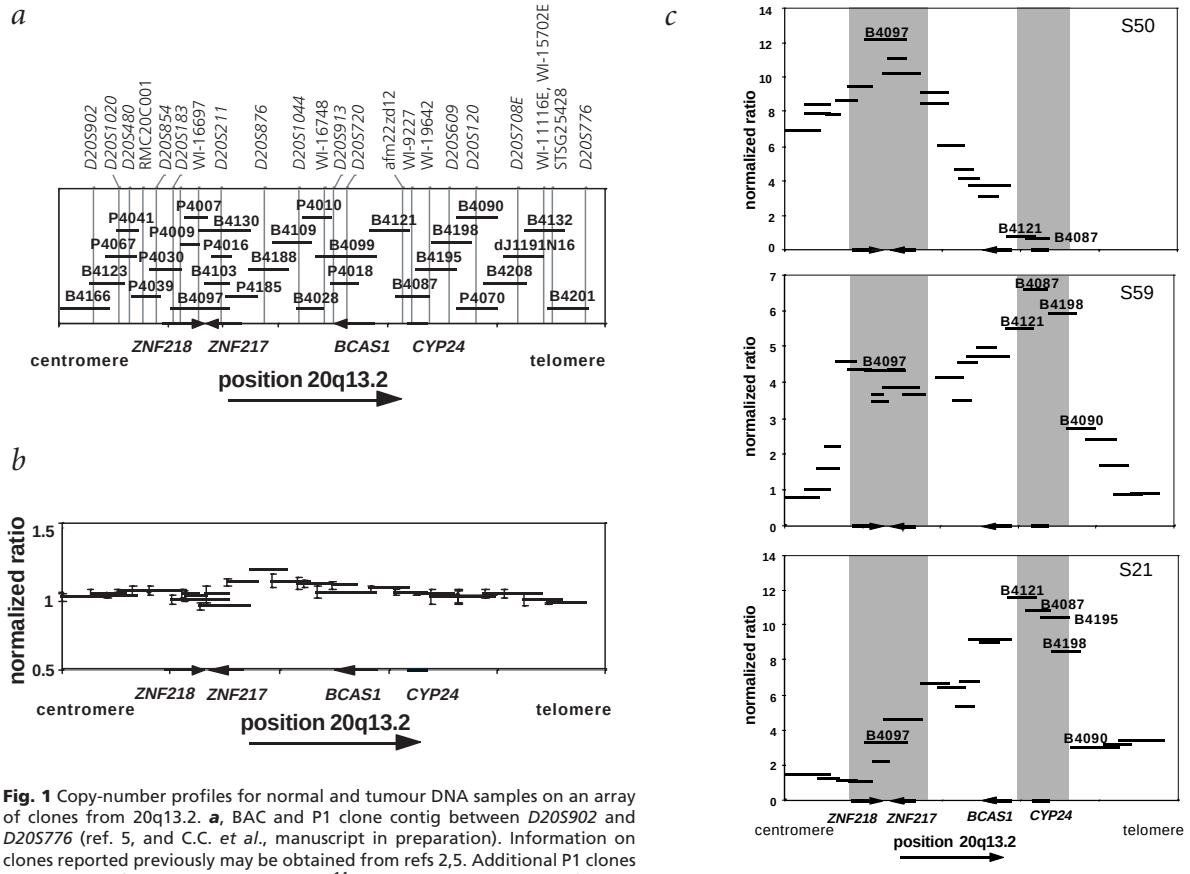

Fig. 1 Copy-number profiles for normal and tumour DNA samples on an array of clones from 20q13.2. ***a***, BAC and P1 clone contig between *D20S902* and *D20S776* (ref. 5, and C.C. *et al*., manuscript in preparation). Information on clones reported previously may be obtained from refs 2,5. Additional P1 clones were selected from the Du Pont B library[14] and included RMC20P4009 (coordinates 35 F9) and RMCP4016 (coordinates 58 B9). Additional BAC clones obtained from the Research Genetics Human BAC libraries[15] included RMC20B4121 (CITB 10 C10), RMC20B4195 (CITB 10 E24), RMC20B4198 (CITB 222 D11), RMC20B4090 (CITB 51 B21), RMC20B4208 (RPCI-11 31D8), RMC20B4132 (CITB 262 J18) and RMC20B4201 (RPCI-11 168E12). The clone dJ1191N16 (RPCI-5 119N16) was obtained from the Sanger Centre. The lengths of the horizontal lines representing the clones are proportional to the clone sizes or STS content. Names are given above the line, omitting the RMC20 prefix. Selected STS markers are shown above the contig and their positions are indicated by vertical lines. The location of *ZNF218*, *ZNF217*, *BCAS1* and *CYP24* is indicated below the contig (arrowheads indicate transcriptional polarity, where known). ***b***, Fluorescence ratios on the chromosome 20q13.2 array targets for three comparisons of a normal genome to itself. Fabrication of arrays, comparative hybridizations, imaging and analysis were carried out as described[2]. Each array target clone is represented by a horizontal bar indicating the location and length of the clone as determined by STS content mapping. The vertical position indicates the measured ratio. The ratios were normalized so that the average for all targets in each hybridization was 1.0. The data points show the mean of the three normalized ratios obtained for each target, and the error bars indicate the standard deviations. The overall standard deviation for the measurements was 0.06. ***c***, Copy-number profiles for three tumours (S50, S59 and S21) showing variation in copy number across the region. Ratios on clones from the contig (*a*) were normalized to the average of the ratios of six clones from chromosome 20p (ref. 2). The data have been plotted as in (*b*). The two regions of peak copy number are shaded. Steep copy-number transitions within RMC20B4198 occurred in two tumours, suggesting that the site may be particularly prone to rearrangement. The average coefficient of variation of the ratios was 14.5% for three separate hybridizations with S21 genomic DNA.

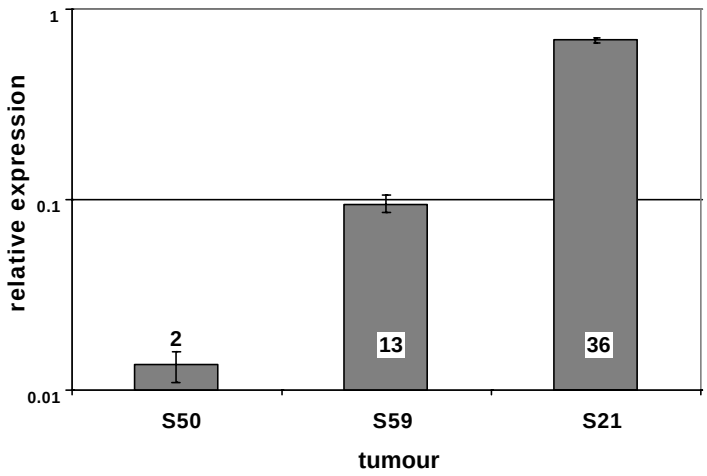

Fig. 2 Relative expression of *CYP24* in breast tumours. Transcript levels of *CYP24* and *VDR* in RNA from frozen breast tumour sections were measured by quantitative RT–PCR. Amplification reactions were performed in triplicate on three different concentrations of each tumour RNA. A 71-bp fragment starting at position 783 in the *CYP24* mRNA sequence (GenBank accession no. NM 000782) was amplified using forward primer 5′–CAAACCGTGGAAGGCCTATC–3′ and reverse primer 5′–AGTCTTCCCCTTCCAGGATCA–3′. A TaqMan probe (5′–FAM (6-carboxy-fluorescein)-ACTACCGCAAAGAAGGCTACGGGCTG–3′ TAMRA (6-carboxy-tetramethyl-rhodamine)) was included in the reaction. An 81-bp fragment starting at position 246 in the *VDR* mRNA sequence (GenBank accession no. NM 000376) was amplified using forward primer 5′–CTTCAGGCGAAGCATGAAGC–3′ and reverse primer 5′–CCTTCATCATGCCGATGTCC–3′. A TaqMan probe (5′–FAM-AAGGCACTATTCACCTGCCCCTTCAA–3′ TAMRA) was included. The data are plotted as $2^{-\Delta N \pm SD}$, where ΔN is the difference in the number of cycles needed for the fluorescence intensity for each gene to reach a threshold value. Copy number of the *CYP24* locus in each tumour is shown and was calculated from the array CGH ratios and the number of copies of 20p, as determined by FISH. Relative levels of *CYP24* were higher in the tumours with amplification at the *CYP24* locus and highest in the most highly amplified tumour, S21.

receptor[6] (VDR), we measured both *CYP24* and *VDR* transcript levels using quantitative PCR. Expression of *CYP24*, normalized with respect to *VDR*, correlates with copy number of *CYP24* in tumours (Fig. 2), further supporting an oncogenic role for *CYP24*.

Our observations indicate that tumours may show peaks in the copy-number profile across an amplified region. Extrapolation from model systems suggests that the peaks contain genes that are selected during the amplification process. Both *CYP24* and *ZNF217* (ref. 5), two genes located at the amplification maxima at 20q13.2, show expression increases consistent with assignment as putative oncogenes. Furthermore, the function of *CYP24* suggests the manner in which the gene contributes to oncogenesis. Thus, the capability to identify peaks in copy-number profiles across regions of recurrent abnormality may be generally useful for localization of oncogenes. Our analysis also indicates that mapping amplicon boundaries alone may provide misleading information on the location of critical genes when there are multiple, closely spaced amplicons. Poor prognosis[12] and aggressive clinical features[12,13] have been associated with increased copy number at the RMC20C001 locus, which maps to the proximal end of the 20q13.2 contig (Fig. 1*a*). As this locus is not centred on either of the regions reported here, additional studies will be required to determine the prevalence and possible specific phenotypic consequences of these more precisely defined amplifications.

Acknowledgements

This work was supported by NIH grants CA80314, CA45919, HD17665 and P50 CA58207, California BCRP grant 2RB-0225 and Vysis.

Donna G. Albertson[1,2], Bauke Ylstra[1], Richard Segraves[2], Colin Collins[1,2], Shanaz H. Dairkee[3], David Kowbel[1], Wen-Lin Kuo[2], Joe W. Gray[2] & Daniel Pinkel[2]

[1]Cancer Research Institute, University of California, San Francisco, Box 0808, San Francisco, California, USA. [2]Cancer Genetics and Breast Oncology Programs, UCSF Cancer Center, University of California, San Francisco, Box 0808, San Francisco, California, USA. [3]Geraldine Brush Cancer Research Institute, California Pacific Medical Center, San Francisco, California, USA. Correspondence should be addressed to D.G.A. (e-mail: albertson@cc.ucsf.edu).

1. Solinas-Toldo, S. *et al. Genes Chromosomes Cancer* **20**, 399–407 (1997).
2. Pinkel, D. *et al. Nature Genet.* **20**, 207–211 (1997).
3. Geschwind, D.H. *et al. Dev. Genet.* **23**, 215–229 (1998).
4. Pollack, J.R. *et al. Nature Genet.* **23**, 41–46 (1999).
5. Collins, C. *et al. Proc. Natl Acad. Sci. USA* **95**, 8703–8708 (1998).
6. Walters, M.R. *Endocr. Rev.* **13**, 719–764 (1992).
7. Albertson, D.G. *Genetics* **134**, 211–219 (1993).
8. Albertson, D.G., Fishpool, R.M. & Birchall, P.S. *Caenorhabditis elegans: Modern Biological Analysis of an Organism* 339–364 (Academic, New York, 1995).
9. Brodeur, G.M. & Hogarty, M.D. *The Genetic Basis of Human Cancer* 161–172 (McGraw-Hill, New York, 1998).
10. John, E.M., Schwartz, G.G., Dreon, D.M. & Koo, J. *Cancer Epidemiol. Biomarkers Prev.* **8**, 399–406 (1999).
11. Alini, M., Marriott, A., Chen, T., Abe, S. & Poole, A.R. *Dev. Biol.* **176**, 124–132 (1996).
12. Tanner, M.M. *et al. Clin. Cancer Res.* **1**, 1455–1461 (1995).
13. Courjal, F. *et al. Cancer Res.* **57**, 4360–4367 (1997).
14. Shepherd, N.S. *et al. Proc. Natl Acad. Sci. USA* **91**, 2629–2633 (1994).
15. Kim, U.J. *et al. Proc. Natl Acad. Sci. USA* **93**, 6297–6301 (1996).