SHORT REPORT

Blood-derived gene-expression profiling in unravelling susceptibility to recessive disease

P Vahteristo, A Kokko, O Saksela, K Aittomäki, L A Aaltonen

J Med Genet 2007;**44**:718–720. doi: 10.1136/jmg.2007.051342

Identification of new disease predisposition genes with chip-based technologies typically requires extensive financial and sample resources. We have recently shown that combining peripheral blood genome and transcriptome (BGT) information in highly selected materials can be a successful low-cost approach to unravelling dominant tumour susceptibility. In this study, we extended our investigations to recessively inherited tumour predisposition, and identified a homozygous germline mutation in the *damage-specific DNA binding protein 2 (DDB2)* gene in a patient with several facial tumours, for which doctors had been unable to provide a diagnosis. Our results provide proof of principle that BGT is a powerful approach for both dominant and recessive genes. In addition to tumour susceptibility, the method may be useful in characterising genetic defects underlying other disease phenotypes.

A 12-year-old boy had been diagnosed with nasal spinocellular carcinoma and soon after with another facial spinocellular carcinoma, two small basalomas, and focal dysplastic and carcinoma *in situ* changes of the facial tissue. The phenotype was suggestive of xeroderma pigmentosum (XP), although less severe than expected. XP is a condition in which the affected patients are hypersensitive to UV irradiation and usually develop several tumours in sun-exposed areas. The phenotype of patients with XP is due to recessively inherited defects in the nucleotide excision repair (NER) pathway, which is responsible for the recognition and removal of multifunctional DNA lesions that arise after UV exposure. Altogether, seven different XP complementation groups are known (A–G), and the underlying genes for each group have been identified (for review of XP and the underlying gene defects, see Cleaver¹).

METHODS

Fibroblastoid cells taken from the patient were sent to the diagnostics laboratory for measurement of unscheduled DNA synthesis (UDS), which is routinely used in XP diagnostics; however, the results were normal. We hence hypothesised that the phenotype might have arisen due to a mutation in an as yet unidentified susceptibility gene. As the patient also had progressive myoclonus epilepsy and a homozygous *cystatin B* mutation, we reasoned that the parents were probably distantly related and that the gene defect underlying the skin tumours was probably also recessively inherited. We hypothesised that performing a BGT analysis in the parents and child could reveal the underlying susceptibility gene, as we had found in our recent study in which dominant low-penetrance germline mutations in the *AIP* gene were identified in pituitary adenoma predisposition.² In BGT, the rationale is to combine genome-wide mapping data with transcriptome-wide expression data to identify aberrantly expressed genes in the genomic areas of interest (fig 1). Statistical analysis of expression arrays

combined with genomic information from single-nucleotide polymorphism arrays is a powerful method to mine the array data. By concentrating on the most relevant regions of the genome and statistically most interesting transcripts, candidate gene lists can be efficiently reduced. Subsequent data mining of the literature and mutation analysis of most relevant targets may enable cost-effective identification of the predisposing gene. Mutations that may lead to abnormal expression include aberrations in regulatory elements, copy-number variations, and truncating mutations leading to nonsense-mediated mRNA decay.

RNA from blood samples taken from the patient and his parents, and from 12 controls available from other projects, were subjected to whole-genome expression arrays (Human_GeneChip_U133_Plus_2.0; Affymetrix, Santa Clara, California, USA). The normalised expression data were filtered using flag calls determined by the Affymetrix detection algorithm: all values receiving the “present” or “marginal” flag were used, whereas values with an “absent” flag were marked as “no data”.

Statistical analysis

Data processing and analyses were performed with GeneSpring V.7.0 software (Agilent Technologies, Santa Clara, California, USA). The *t* test was used to compare the expression between the parents and controls, and *p* < 0.05 was considered statistically significant. All the expression data have been deposited in the GEO database (accession number GSE6351).

RESULTS

Altogether, 42 probe sets were “absent” in the patient’s sample but “present” in all the control samples; expression of 10 of these probe sets in the parents’ samples was significantly reduced, as expected with heterozygosity (table 1). Interestingly, one of the 10 genes was *damage-specific DNA binding protein 2 (DDB2)*, which was considered an attractive candidate for the observed phenotype. A literature search revealed that <10 biallelic *DDB2* mutations had already been described in patients with XP type E (XPE).^{3–6} Subsequent sequencing revealed a nonsense mutation 574C→T (R192X) in *DDB2*, with the parents being heterozygous and the patient homozygous for the mutation. *DDB2* encodes for the p48 protein, which together with p127 forms a heterodimeric complex responsible for the recognition of damaged DNA.⁵ In normal human cells, expression of *DDB2* is induced by UV light through p53-dependent transcription.⁷ Cells with non-functional *DDB2* are therefore unable to recognise the various forms of DNA damage caused by UV irradiation.

Strikingly, samples from only this single affected patient and his healthy carrier parents were sufficient to reveal the gene

Abbreviations: BGT, blood genome and transcriptome; *DDB2*, *damage-specific DNA binding protein 2*; NER, nucleotide excision repair; UDS, unscheduled DNA synthesis; XP, xeroderma pigmentosum

Figure 1 Method for peripheral blood genome and transcriptome (BGT) analysis.

defect. We had planned to perform a chip-based genomic homozygosity analysis on the patient’s sample to exclude aberrantly expressed genes residing in heterozygous regions, but this was unnecessary. Importantly, however, it is clear that in less favourable circumstances, BGT would have provided very significant added value by reducing the number of candidate genes to a fraction of the original possibilities.

DISCUSSION

We report a case of XP in which all routine diagnostic tests were negative, compatible with previous experience reported in XPE.^{6,8} Because the correct diagnosis and understanding of the condition is essential for motivating patients to permanently take the strictest possible UV avoidance measures, missing the diagnosis in clinically less severely affected patients can have serious future consequences for them.

As proof of principle, our results highlight the power of new array-based technologies in resolving origins of human disease susceptibility that arise due to mutations that alter the expression of the predisposition genes. To date, we have shown the applicability of the method for both dominant² and recessive conditions; in this study, the mode of inheritance was recessive, and samples from the single patient and his healthy parents were sufficient for gene identification. To

examine the number of samples required in a dominantly inherited disease, we re-analysed the data that led to the identification of *AIP* as a new pituitary tumour predisposition gene.² Originally the gene expression was compared between nine cases and seven controls. Based on our simulations, it became evident that the analysis of as few as any three cases versus the seven controls would have been sufficient to reveal the underlying gene defect (data not shown). The exact specimen numbers necessary to detect a difference are of course subject to parameters such as numbers of genes in the genomic areas of interest and the degree of altered expression caused by the genetic defect. However, it is obvious that success is possible with remarkably few cases. Therefore, BGT enables the effective use of small amounts of unique materials, leading to enhanced and cost-effective identification of disease-predisposition genes. The approach may also assist in resolving the genetic defect in disorders with extreme locus heterogeneity. For example, in the recessively inherited Bardet–Biedl syndrome there are already 12 known predisposing genes, which together account for approximately 75% of the cases.⁹ As most mutations are truncating and affect the expression of the corresponding protein, an expression-based approach might be useful in pinpointing the gene in question, either known or novel genes. When using BGT, especially for diagnosis, it is important to bear in mind that it can only detect mutations that have an effect on expression. Therefore, although it may point in the right direction in the case of a positive result, little can be concluded if the result for a particular gene is negative—that is, if no change of expression is detected.

We have experienced failure of BGT twice in research: in an effort to identify a gene for the paraganglioma 2 locus on chromosome 11 (though the effort is not yet fully completed),¹⁰ and in identifying a putative colorectal adenoma-susceptibility gene on chromosome 15.¹¹ The potential reasons for failure of BGT have been discussed above, and it is difficult to determine the reason in any particular case until other approaches have identified the respective culprit defect. Before using BGT to identification of *AIP*, we had tested the power of the method to identify the *fumarase* (chromosome 1) and *MLH1* (chromosome 3) tumour-predisposition genes, and found that in both cases BGT would have been useful in gene identification, encouraging us to use it in the *AIP* study.² The success of BGT partly depends on the expression of the culprit gene in lymphocytes, but the phenotypes examined to date provide few clues about this issue. On the one hand, a condition affecting multiple organ systems such as an *MLH1* defect, seems to be a promising target for BGT, but we have recently derived promising data in an ongoing effort to understand the genetic basis of tylosis with oesophageal cancer syndrome mapping to chromosome 17,¹² even though the tylosis phenotype is limited to stratified squamous epithelia.

Table 1 Probe sets that were absent in the index patient and showed reduced expression (p<0.05) in the parents compared with controls

Probe set ID	Gene symbol	Gene name/description	Fold change*	p Value†
236769_at	–	Unknown	0.40	<0.001
203409_at	<i>DDB2</i>	Damage-specific DNA-binding protein 2	0.56	<0.001
236436_at	–	Unknown	0.70	<0.001
238365_s_at	–	Unknown	0.65	<0.01
236610_at	–	Unknown	0.60	<0.01
221631_at	<i>CACNA11</i>	Voltage-dependent T-type calcium channel subunit α -11	0.75	0.01
232264_at	–	Unknown	0.43	0.03
219040_at	<i>CORO7</i>	Coronin 7	0.80	0.03
228040_at	<i>MGC21881</i>	MGC21881 protein	0.76	0.03
203691_at	<i>PI3</i>	Protease inhibitor WAP3	0.42	0.04

*Parents/controls.

†Parents compared with controls, t test; not corrected for multiple testing.

Key points

- BGT can be a successful low-cost approach in unravelling tumour susceptibility in both dominant and recessive conditions.
- BGT can be useful in characterising phenotypes with an unidentified genetic component and perhaps also when tackling challenging diagnostic cases.
- In addition to tumour susceptibility, the method may be useful in characterising genetic defects underlying other disease phenotypes.

Overall, the feasibility of using blood-derived RNA in expression arrays, the decreasing costs of chip-based technologies, and the limited number of samples required provide a platform for wider use of the BGT approach in characterising phenotypes with an unidentified genetic component and perhaps also when tackling challenging diagnostic cases. Our experience suggests that the strategy we propose is very cost-effective and is applicable to both dominant and recessive conditions. In addition to tumour susceptibility, BGT could potentially be used to characterise genetic defects underlying other disease phenotypes.

ACKNOWLEDGEMENTS

We wish to thank the family members participating in the study, and Iina Vuoristo and Sini Marttinen for excellent technical assistance.

Authors' affiliations

P Vahteristo, A Kokko, L A Aaltonen, Department of Medical Genetics, University of Helsinki, Helsinki, Finland
O Saksela, Department of Dermatology, Helsinki University Central Hospital (HUCH), Helsinki, Finland
K Aittomäki, Department of Clinical Genetics, HUCH, Helsinki, Finland

Competing interests: None.

Funding: This work has been financially supported by the Academy of Finland (Center of Excellence in Translational Genome-Scale Biology 2006–2010, grant 212901 for PV) and Finnish Cultural Foundation.

Ethics approval: Patient information and samples were obtained with full informed consent. The study was approved by the appropriate ethics review committee of the Hospital District of Helsinki and Uusimaa.

Correspondence to: Lauri A Aaltonen, Department of Medical Genetics, Biomedicum Helsinki, PO Box 63, 00014 University of Helsinki, Finland; lauri.aaltonen@helsinki.fi

Received 3 May 2007

Revised 24 June 2007

Accepted 29 June 2007

Published Online First 27 July 2007

REFERENCES

- 1 **Cleaver JE**. Cancer in xeroderma pigmentosum and related disorders of DNA repair. *Nat Rev Cancer* 2005;**5**:564–73.
- 2 **Vierimaa O**, Georgitsi M, Lehtonen R, Vahteristo P, Kokko A, Raitila A, Tuppurainen K, Ebeling TML, Salmela PI, Paschke R, Gündogdu S, De Menis E, Mäkinen MJ, Launonen V, Karhu A, Aaltonen LA. Pituitary adenoma predisposition caused by germline mutations in the AIP gene. *Science* 2006;**312**:1228–30.
- 3 **Nichols AF**, Ong P, Linn S. Mutations specific to the xeroderma pigmentosum group E Ddb – phenotype. *J Biol Chem* 1996;**271**:24317–20.
- 4 **Itoh T**, Mori T, Ohkubo H, Yamaizumi M. A newly identified patient with clinical xeroderma pigmentosum phenotype has a non-sense mutation in the DDB2 gene and incomplete repair in (6-4) photoproducts. *J Invest Dermatol* 1999;**113**:251–7.
- 5 **Nichols AF**, Itoh T, Graham JA, Liu W, Yamaizumi M, Linn S. Human damage-specific DNA-binding protein p48. Characterization of XPE mutations and regulation following UV irradiation. *J Biol Chem* 2000;**275**:21422–8.
- 6 **Rapic-Otrin V**, Navazza V, Nardo T, Botta E, McLenigan M, Bisi DC, Levine AS, Stefanini M. True XP group E patients have a defective UV-damaged DNA binding protein complex and mutations in DDB2 which reveal the functional domains of its p48 product. *Hum Mol Genet* 2003;**12**:1507–22.
- 7 **Tan T**, Chu G. p53 binds and activates the xeroderma pigmentosum DDB2 gene in humans but not mice. *Mol Cell Biol* 2002;**22**:3247–54.
- 8 **Itoh T**, Linn S, Ono T, Yamaizumi M. Reinvestigation of the classification of five cell strains of xeroderma pigmentosum group E with reclassification of three of them. *J Invest Dermatol* 2000;**114**:1022–9.
- 9 **Stoetzel C**, Muller J, Laurier V, Davis EE, Zaghloul NA, Vicaire S, Jacquelin C, Plewniak F, Leitch CC, Sarda P, Hamel C, de Ravel TJ, Lewis RA, Friederich E, Thibault C, Danse JM, Verloes A, Bonneau D, Katsanis N, Poch O, Mandel JL, Dollfus H. Identification of a novel BBS gene (BBS12) highlights the major role of a vertebrate-specific branch of chaperonin-related proteins in Bardet-Biedl syndrome. *Am J Hum Genet* 2007;**80**:1–11.
- 10 **Mariman EC**, van Beersum SE, Cremers CW, Struycken PM, Ropers HH. Fine mapping of a putatively imprinted gene for familial non-chromaffin paragangliomas to chromosome 11q13.1: evidence for genetic heterogeneity. *Hum Genet* 1995;**95**:56–62.
- 11 **Jaeger EE**, Woodford-Richens KL, Lockett M, Rowan AJ, Sawyer EJ, Heinemann K, Rozen P, Murday VA, Whitelaw SC, Ginsberg A, Atkin WS, Lynch HT, Southey MC, Debinski H, Eng C, Bodmer WF, Talbot IC, Hodgson SV, Thomas HJ, Tomlinson IP. An ancestral Ashkenazi haplotype at the HMP5/CRAC1 locus on 15q13-q14 is associated with hereditary mixed polyposis syndrome. *Am J Hum Genet* 2003;**72**:1261–7.
- 12 **McDonald FE**, Liloglou T, Xinarianos G, Hill L, Rowbottom L, Langan JE, Ellis A, Shaw JM, Field JK, Risk JM. Down-regulation of the cytoglobin gene, located on 17q25, in tylosis with oesophageal cancer (TOC): evidence for trans-allele repression. *Hum Mol Genet* 2006;**15**:1271–7.