Published in final edited form as:

Clin Cancer Res. 2010 November 1; 16(21): 5211–5221. doi:10.1158/1078-0432.CCR-10-1521.

Patterns of expression of DNA repair genes and relapse from melanoma

Rosalyn Jewell⁺, Caroline Conway⁺, Angana Mitra, Juliette Randerson-Moor, Samira Lobo, Jérémie Nsengimana, Mark Harland, Maria Marples^{*}, Sara Edward^{*}, Martin Cook^{**}, Barry Powell^{***}, Andy Boon⁺, Floor de Kort^{****}, Katharine A Parker^{****}, Ian A. Cree^{*****}, Jennifer H. Barrett, Margaret A. Knowles, D. Timothy Bishop, and Julia Newton-Bishop

Section of Epidemiology and Biostatistics, Leeds Institute of Molecular Medicine, St James's University Hospital, Beckett Street, Leeds

^{*}Leeds Teaching Hospital Trust, UK ^{**}Royal Surrey County Hospitals NHS Trust, Surrey, UK ^{***}St George's Hospital Melanoma Unit, Blackshaw Road, Tooting, London ^{****}ServiceXS B.V. Leiden, Netherlands ^{*****}Translational Oncology Research Centre, Queen Alexandra Hospital, Portsmouth

Abstract

Purpose—To use gene expression profiling of formalin-fixed primary melanoma samples to detect expression patterns that are predictive of relapse and response to chemotherapy.

Experimental design—Gene expression profiles were identified in samples from two studies (472 tumors). Gene expression data for 502 cancer-related genes from these studies were combined for analysis.

Results—Increased expression of DNA repair genes most strongly predicted relapse, and was associated with thicker tumors. Increased expression of *RAD51* was the most predictive of relapse-free survival (RFS) in unadjusted analysis (HR 2.98, $p=8.80 \times 10^{-6}$). *RAD52* (HR 4.73, $p=0.0004$) and *TOP2A* (HR 3.06, $p=0.009$) were independent predictors of RFS in multivariable analysis. These associations persisted when the analysis was further adjusted for demographic and histological features of prognostic importance (*RAD52* $p=0.01$; *TOP2A* $p=0.02$). Using principal component analysis, expression of DNA repair genes was summarized into one variable. Genes whose expression correlated with this variable were predominantly associated with the cell cycle and DNA repair. In forty-two patients treated with chemotherapy, DNA repair gene expression was greater in tumors from patients who progressed on treatment. Further data supportive of a role for increased expression of DNA repair genes as predictive biomarkers are reported, which were generated using multiplex polymerase chain reaction (PCR).

Conclusions—Over-expression of DNA repair genes (predominantly those involved in double-strand break repair) was associated with relapse. These data support the hypothesis that melanoma

Corresponding author Dr. Rosalyn Jewell Clinical Research Fellow Section of Epidemiology and Biostatistics, Leeds Institute of Molecular Medicine, St James's Hospital, Beckett Street, Leeds. LS9 7TF Tel +44 113 2065037 Fax +44 113 2340183
r.a.jewell@leeds.ac.uk

⁺These authors contributed equally to this work

Previous presentations: Part of this work was presented in oral presentations at the European Organization for Research and Treatment of Cancer (EORTC) Melanoma Group Meeting, Belgium, March 2009; the 7th World Congress on Melanoma, Vienna, May 2009 and the Melanoma Study Group Meeting, Cambridge, June 2009.

Conflicts of interest: none

progression requires maintenance of genetic stability and give insight into mechanisms of melanoma drug resistance and potential therapies.

Keywords

formalin-fixed; tumor; prognosis; biomarker; DASL; DNA repair

Background

The established predictors of outcome for melanoma patients relate to the histological characteristics of the primary tumor (Breslow thickness, presence of ulceration(1), mitotic rate(2)), tumor site, sex(2) and age(3). These histological characteristics are used to estimate prognosis as part of the AJCC staging system(1) and in various algorithms(2, 3), but much of the variance in survival remains unexplained. In order to identify prognostic and predictive biomarkers, and to better establish the biological pathways of relevance, molecular studies of primary melanomas are necessary.

Genomic and transcriptomic studies in melanoma tumors have hitherto been few, using limited numbers of samples due to the small size of melanoma primaries. A number of studies using cryopreserved tissue have identified gene signatures or altered expression in groups of genes within a biological pathway that are predictive of survival or progression of a tumor(4-7). One such study by Kauffmann *et al.* demonstrated, in sixty cryopreserved melanomas, that over-expression of DNA repair genes was associated with metastasis, a finding validated in a further seventeen tumors(7).

We have recently reported the use of Illumina's DASL (cDNA mediated annealing, selection, extension and ligation) assay to discover prognostic biomarkers in formalin-fixed paraffin-embedded primary melanomas(8). We obtained gene expression detection from 76% of unselected primary tumor blocks and identified increased expression of osteopontin as being the single change most predictive of relapse in samples from a cohort of melanoma patients from the UK(8). Here we report that, among the 502 cancer-related genes studied, the group of genes whose expression was most strongly predictive of relapse overall, in a larger pooled data set, were a number of DNA repair genes or genes involved in this process (predominantly in double-strand break (DSB) repair). We went on to explore the possibility that increased expression of these genes might also be useful predictive biomarkers of response to chemotherapy.

Methods

Patient Samples

Formalin-fixed paraffin-embedded cutaneous primary tumors were identified from two study sets as previously described(8) (Table 1s). In Study 1, the "Leeds Melanoma Cohort Study"(9), population ascertained melanoma patients were recruited to a cohort in the period from 2000 to 2006. 254 blocks identified from participants within the cohort with tumors thicker than 0.75 mm with the longest follow up were used in this study. In Study 2, 218 blocks identified from participants with the longest follow-up in a study designed to identify predictors of sentinel node positivity and relapse, were sampled. Both studies were approved by UK national ethics committees (MREC and PIAG). Here we report the combined gene expression data analysis from these studies.

Sample Preparation

Primary tumor blocks were identified, and a hematoxylin and eosin (H&E) stained slide was examined to identify the deepest part of the tumor, containing the lowest admixture of

inflammatory or stromal cells. This area was marked using a fine tipped permanent marker and a 0.8 mm tissue microarray (TMA) needle was used to sample the tumor block horizontally. This technique produces sufficient RNA yields from the deepest parts of tumors in a study using hundreds of samples, while preserving the block for future pathology(8).

RNA extraction and quality control

Tissue cores were de-waxed using xylene, which was removed using two changes of absolute ethanol. RNA was extracted from tissue cores using the High Pure paraffin RNA kit (Roche Diagnostics Ltd, Burgess Hill, UK) according to the manufacturer's protocol and eluted in 26 µl nuclease free water. Details of RNA quality control procedures have been described previously(8).

Gene expression profiling with the DASL assay

The DASL assay was performed by Service XS, Leiden using the Human Cancer Panel (Illumina Inc., CA, USA). The Cancer Panel includes 1536 unique sequence specific probes targeting 502 genes (listed in supplementary information). Each gene is targeted in three locations by three separate probe pairs designed by a proprietary algorithm(10).

Data pre-processing

The data were normalized using Beadstudio software (Illumina, USA) before exporting to STATA version 10 for statistical analyses. The normalization methods used were background correction, cubic spline smoothing(11) and plate scaling. Normalization was conducted relative to a synthetic reference array, which was created in each study by averaging all melanoma samples as described previously(8). Each study was normalized separately prior to the data sets being combined for analysis.

Statistical methodology

In the DASL assay, the number of genes detected in each sample (probe signal significantly greater than average signal from negative controls with $p < 0.05$) was used as a measure of the quality of the results. Samples with fewer than 250 detected genes were classified as 'failed' and excluded from further analysis.

Mean gene expression was used for replicate samples. Fold change calculations were performed using normalized data. Differential gene expression, survival analyses and chemotherapy analyses were performed using log-transformed normalized data (log2). Within the sample sets, expression of each gene was compared with histological features of interest or response status to chemotherapy using linear regression. Chemotherapy response was defined as stable disease, mixed response, partial response or regression of disease determined radiologically (by CT scan) following 2 or 3 cycles of chemotherapy.

Survival analysis was performed using a Cox proportional hazards model to estimate hazard ratios and 95% confidence intervals for each gene. Significance values were ranked to identify the genes most differentially expressed between groups of interest. Relapse free survival (RFS) was defined as the period between diagnosis and date of first relapse at any site. To identify DNA repair genes that were independently predictive of survival, a multivariable Cox proportional hazards model was generated using all DNA repair genes significantly associated with Breslow thickness, mitotic rate or survival in the previous analyses.

Ingenuity Pathway Analysis (IPA, Ingenuity Systems, Redwood City, California) was used to identify the pathways and networks of genes co-regulated with the most significant DNA

repair genes. To achieve this goal, a three-stage approach was adopted: (1) To summarize into one variable the expressions of the most significant DNA repair genes, principal components analysis was applied to the pooled data set adjusting for study. (2) A correlation coefficient was calculated between the first principal component (summary variable) and each of the genes on the cancer panel. (3) The IPA was interrogated to find pathways and networks involving genes significantly correlated with the principal component at a level of $p < 10^{-10}$.

For analyses identifying genes predictive of survival and genes associated with histological features, the Bonferroni method was used to correct for multiple testing(12); therefore the significance level was set at 0.0001 for these analyses. For chemotherapy and survival analysis assessing the expression of single genes in each test, the significance level for highlighting results of interest was set at 0.05. All analyses were adjusted for the study from which the patients were recruited and survival analysis was further adjusted for whether the patient had undergone a sentinel node biopsy and the effect of the biopsy result (SNB status). Analyses were further adjusted for demographic and histological factors of prognostic importance in melanoma.

Target validation by quantitative Real-Time reverse transcription polymerase chain reaction

DNA repair genes identified as significantly associated with relapse were further investigated by quantitative Real-Time reverse transcription polymerase chain reaction (qRT-PCR) on samples from Study 1 using Taqman® Gene Expression Assays (Applied Biosystems, Warrington, UK). The comparative Ct method(13) was used to calculate relative fold changes in gene expression (supplementary information). To identify genes associated with chemotherapy response, a selection of samples from patients treated with chemotherapy were analyzed using a customized Taqman Array microfluidic qRT-PCR card (Chemo-sensitivity Gene Expression Assay, CGEA-1, CanTech Ltd, Portsmouth, UK) containing 92 genes known or hypothesized to be involved in cytotoxic resistance or sensitivity, as reported previously(14). The comparative Ct method(13) was used to calculate relative fold changes in gene expression between non-responders and responders to chemotherapy (supplementary information).

Results

Patient and tumor characteristics of the two samples sets were similar and are representative of the larger studies that the test sets were derived from as previously described(8) (Table 1s).

RNA yields from tumor samples and performance using the DASL assay

Adequate RNA yields were obtained from 361/472 (76%) of blocks as described previously(8). Four hundred and twenty three RNA samples including replicate samples (359 unique samples and 64 replicate samples) were supplied to ServiceXS. Fewer than 250 genes were detected in 6 (1.4%) samples (failed samples). The failure rate was 2.1% in Study 1 and 0.9% in Study 2(8).

Genes predictive of relapse

Genes whose expression levels in tumors were most strongly related to RFS are presented in Table 1 when analysis is adjusted for study type and SNB status only. The predominant group of genes differentially expressed were DNA repair genes or genes associated with these processes. Within this group, over-expression of *RAD51*, *RAD52* and *TOP2A* were most predictive of poor RFS (Table 2) with hazard ratios of 2.98 ($p = 8.80 \times 10^{-6}$), 4.49

($p=0.00002$) and 3.84 ($p=0.00009$) respectively for a doubling of expression levels. Over-expression of these genes continued to be predictive of RFS when the analysis was adjusted for host variables known to predict relapse (age, sex and tumor site), and when adjusted additionally for Breslow thickness, mitotic rate (number/mm²) and the presence of tumor ulceration (Table 2). Expression of *RAD51* was 1.22 times greater in tumors from patients who relapsed versus those that did not; the fold changes between tumors from relapsers and non-relapsers for *RAD52* and *TOP2A* were 1.16 and 1.12 respectively (Table 1). *RAD54B*, *RAD52*, *TOP2A* and *RAD51* were also over-expressed in tumours from patients who died versus surviving patients (fold changes of 1.15, 1.11, 1.09 and 1.10, respectively).

Independent predictors of RFS

Over-expression of *RAD52* and *TOP2A* were independent predictors of poor RFS in a model with the other DNA repair genes identified (listed in Table 3) (multivariable model Table 2 and Figure 1). In analysis adjusted for study and SNB status only, hazard ratios (HR) for *RAD52* and *TOP2A* were 4.73 ($p=0.0004$) and 3.06 ($p=0.009$), respectively, for a doubling of levels. When analyzed by quartiles of *RAD52* and *TOP2A* gene expression, HRs increased for each quartile of gene expression: using the lowest quartile as baseline, *RAD52* 25-50% expression HR 1.57 (95% CI 0.80-3.09, $p=0.19$), 50-75% expression HR 1.79 (95% CI 0.95-3.37, $p=0.07$), over 75% expression HR 2.90 (95% CI 1.57-5.36, $p=0.001$); *TOP2A* 25-50% expression HR 2.82 (95% CI 1.46-5.47, $p=0.002$), 50-75% expression HR 2.67 (95% CI 1.34-5.29, $p=0.005$) and over 75% expression HR 3.59 (95% CI 1.78-7.24, $p<0.0001$) (Figure 1). Both *RAD52* and *TOP2A* continued to have a significant independent predictive influence on RFS when analyses were adjusted for host variables (age, sex and tumor site) and further adjusted for Breslow thickness, mitotic rate and ulceration of the tumor (Table 2).

Gene expression in tumors with poor prognostic histopathological features

Genes involved in DNA repair were also relatively over-expressed in tumors with greater Breslow thickness and mitotic rate (Table 3).

Target validation by quantitative Real-Time PCR

To seek corroborative data for the DASL assay results, qRT-PCR was performed on study 1 RNA samples for four DNA repair genes identified as associated with relapse. Increased expression of *RAD51* (fold change 1.11), *RAD54B* (fold change 1.08) and *TOP2A* (fold change 1.37) was detected with qRT-PCR in association with reduced RFS in samples from study 1 compared with fold changes of 1.22, 1.18 and 1.12 respectively in the merged DASL assay analysis. We were unable to validate the association between *RAD52* expression and survival using qRT-PCR in samples from study 1. There was considerable variation in DASL gene expression results for *RAD52* across the three probes, with results from one probe correlating better with qRT-PCR results than the other two. As DASL gene expression data is calculated from the mean of three probes, this may explain the discrepancy between DASL and qRT-PCR results.

Co-expression of genes with DNA repair genes

Expression of the majority of DNA repair genes identified as being associated with reduced survival or poor prognostic histological features correlated well with each other (Table 2s). An exception was *RAD52*, with expression levels that did not correlate with the levels of *RAD54L*, *BRCA2* and *TOP2A*.

Principal components analysis of the top 10 significant DNA repair genes (listed in Table 3) identified one component which explained 40% of total variance in DNA repair gene

expression whilst each of the other components explained 10% or less (study adjusted). Principal component 1 was also the only one that correlated with each of the 10 genes from which it was generated (correlation coefficient ranging from 0.40 to 0.74). Association analysis between this component and RFS, Breslow thickness and mitotic rate showed that the component was predictive of these phenotypes when adjusted for study, sex, age and tumor site ($p=1.1\times 10^{-7}$ for Breslow thickness, $p=1.1\times 10^{-10}$ for mitotic rate and $p=3.1\times 10^{-5}$ for RFS). None of the other components was predictive of these phenotypes. In Table 4 we report the list of genes correlated with the first principal component at a level of $p < 10^{-10}$. A large number of these genes are associated with cell cycle control (e.g. *CDC2*, *CCNA2*) or DNA repair (e.g. *PCNA*, *BLM*, *XRCC2*).

The gene network in supplementary Figure 1s built using IPA contains 35 genes among which 22 are correlated with the principal component of DNA repair genes at a significance level of $p<10^{-10}$. The major cellular or molecular functions making up this network are cell cycle and cell death, confirming the strong association between these processes and DNA repair.

Genes predictive of chemotherapy response

Forty-two patients had received chemotherapy (11.9%) so we looked at response to chemotherapy in these individuals as a pilot study. Of these patients, tumor response following chemotherapy had been recorded for 36 patients. The majority of patients received dacarbazine (DTIC) with 3 patients receiving treosulfan as first line chemotherapy. Six patients (17%) responded to the chemotherapy (DTIC). *RAD51* and *TOP2A* had significantly higher expression in tumors from non-responders compared to responders (*RAD51* fold change 1.66, $p=0.01$; *TOP2A* fold change 1.43, $p=0.03$). *BRCA1* was also over-expressed in tumors from non-responders (fold change 1.36), but this difference barely reached statistical significance ($p=0.05$).

Thirty-one of these samples (5 from responders) were additionally sent for qRT-PCR analysis using the CGEA-1 array in Portsmouth. Genes for which there were less than 10 results or less than 2 results for responders were excluded from further analyses. All of the 17 genes involved in DNA repair on the array were over-expressed in tumors from patients who did not respond to chemotherapy (Table 5). These genes included *RAD51* (fold change 2.13), *TOP2A* (fold change 2.53) and *BRCA1* (fold change 1.77) as well as genes involved in nucleotide excision repair (e.g. *ERCC1*, *XPA*), base excision repair (e.g. *XRCC1*), removal of damaged DNA bases (*MGMT*), mismatch repair (e.g. *MLH1*, *MSH2*) and DNA non-homologous end joining (e.g. *XRCC5*, *XRCC6*) (14, 15). Genes encoding proteins associated with apoptosis, cellular proliferation and membrane transport molecules were differentially expressed in tumors which did not respond to chemotherapy (Table 5) (14, 16). As a result of the small numbers of samples included in these analyses, the power was too low to achieve statistical significance.

Discussion

Cutaneous melanoma carries a good prognosis for the majority of patients, but approximately 20% overall die of the disease. Response rates to chemotherapeutic regimens are poor at around 12 to 15% (17) and there are no means of identifying patients who might benefit from treatment. Poor progress in identification of prognostic and predictive biomarkers has been in part a result of the fact that primary melanomas are physically small, and pathologists are therefore reluctant to cryopreserve tumor. Using the Illumina DASL assay Cancer Panel, we have produced gene expression profiles using formalin-fixed material which have prognostic significance (8), giving hope for larger scale biomarker studies in the future.

In this large study, over-expression of DNA repair genes was associated with reduced RFS, greater Breslow thickness and higher tumor mitotic rate. Our results are in agreement with the findings of Kauffmann *et al.* who showed that over-expression of genes involved in recovery of stalled replication forks was associated with metastasis in cryopreserved tumors(7). It has been hypothesized that in order for a melanoma cell to continue to divide and give rise to metastases, the cell needs to maintain genomic integrity and therefore those tumors with up-regulated genes associated in DNA repair processes(18) prove to be more aggressive. Over-expression of DNA repair genes in melanoma tumors may help to explain why melanoma tumors are so resistant to treatment with chemotherapy and radiotherapy(18), a hypothesis which gains some support from our observations in a relatively small number of chemotherapy treated patients.

Many of the genes we have identified as associated with relapse in these analyses are involved in DSB repair via homologous recombination. *RAD51* is the central protein involved in the DSB process and is over-expressed in many tumors(19). *RAD52*, *RAD54*, *BRCA1* and *BRCA2* are required for DSB repair(20) with *RAD52* also involved in single strand annealing independently of *RAD51*(21). *CHEK1* is involved in cell-cycle control; it encodes a protein kinase that prevents progress of the cell from G2 to M phase when activated by DNA damage, such as DSBs and stalled DNA replication forks(7, 15). *MSH2* and *MSH6* are mismatch repair genes that code for proteins that repair errors made in base pairing by DNA polymerases during replication(7, 22). *TOP2A* induces transient DNA breaks to allow changes in DNA topology during DNA replication(23). Expression of *TOP2A* closely reflects the proliferative activity of cells(23).

Formation of DSBs and other DNA damage occurs during replication of cells, so it is unsurprising that, in addition to identifying over-expression of DNA repair genes associated with survival, we have also identified over-expression of genes associated with the cell cycle and cell proliferation in tumors from patients who relapse, and in patients with thicker tumors and with higher mitotic rates. Using principal component analysis of genes associated with RFS, Breslow thickness and mitotic rate, we have shown that many of the genes with expression levels that correlate with the first principal component are involved in DNA repair processes and the cell cycle, indicating that over-expression of genes involved in these pathways is common in tumors with poorer prognosis, perhaps reflecting increased proliferation in these cells. *RAD52* is an exception, as expression levels were not correlated with expression of other genes involved in DSB repair such as *RAD54L* and *BRCA2*. This may reflect the additional function of *RAD52* in repair by single strand annealing(21).

Previous reports have identified the association between *TOP2A* expression and aggressive tumor features in prostate(24), hepatocellular(25) and colorectal cancers(26). Recent work has confirmed that high *TOP2A* gene expression is associated with shorter metastasis-free interval in breast cancer(27). *TOP2A* over-expression in breast tumors correlates with response to anthracycline-based chemotherapy regimens(27, 28), but has also been linked to resistance to chemotherapy regimens in hepatocellular carcinoma (doxorubicin resistance)(25) and colorectal tumors (irinotecan and etoposide resistance)(26).

DTIC remains the standard first line chemotherapy for metastatic melanoma(29) internationally. DTIC is an alkylating agent that attaches alkyl groups to DNA at the O⁶-position of guanine(30), causing inter- and intra-strand cross-links and mutation during cell replication. Finally, DSBs develop in DNA causing cell death(30). The critical toxic lesion induced by DTIC is the addition of alkyl groups, and accordingly the activity of DNA repair genes such as O⁶-methylguanine-DNA methyltransferase (*MGMT*), which removes alkyl groups from DNA, is associated with resistance to alkylating agents in glioblastoma tumors(31) and in melanoma cell lines(32). A limitation of our study is that the 502 genes

assessed on the DASL assay did not include *MGMT*, however in the smaller number of samples assessed using the CGEA-1 array there was increased expression of *MGMT* in tumors unresponsive to chemotherapy which supports these previous findings. The finding that over-expression of *RAD51*, *TOP2A* and *BRCA1* was associated with poor response to chemotherapy is intriguing, suggesting for the first time that genes involved in DSB repair may also be associated with resistance to chemotherapy of melanoma. As DSBs represent the final common pathway of DTIC action in cancer cells(30), this association is biologically plausible. Using the CGEA-1 array, there was increased expression of genes involved in base-excision repair, nucleotide excision repair and DNA non-homologous end joining(15) in tumors unresponsive to chemotherapy. In contrast, genes encoding proteins involved in apoptosis, cellular proliferation and drug pumps exhibited both increased and decreased expression. This observation suggests that increased expression of a number of DNA repair pathways is associated with resistance to chemotherapy in melanoma, with some (limited) evidence of additional mechanisms of chemotherapy resistance involving more complex interactions between cellular pathways. Gene expression studies assessing multiple genes in melanoma tumors have previously been small, and not designed to identify predictive markers; however these current data were derived from a small sample, and the observed effect only reached statistical significance for *RAD51* and *TOP2A*, so these are preliminary observations only. The relationship between higher expression of DNA repair and failure to respond to DTIC may also simply reflect the fact that these tumors are more aggressive, with a higher proliferative rate, rather than a specific relationship with the drug used.

In summary, in this large scale assessment of gene expression profiles in formalin-fixed paraffin-embedded primary melanomas we have identified over-expression of DNA repair genes as being associated with shorter RFS and poor prognostic histopathological features. This validates the findings of Kauffmann *et al.*(7) in a much larger study suggesting that DNA repair pathways are essential biological pathways involved in progression of melanoma and perhaps treatment responses. Although there is hope of new therapeutic agents for the treatment of melanoma it is very likely that a role for DTIC will remain and the identification of predictive biomarkers will remain of crucial importance.

Statement of Translational Relevance

This article presents gene expression results from melanoma primary tumors in a study utilizing formalin-fixed tissue. This is the largest gene expression study to date in melanoma and has identified prognostic and potential predictive markers. Over-expression of DNA repair genes was shown to be associated with reduced relapse-free survival, thicker tumors and tumors with higher mitotic rate. Preliminary data are also reported suggesting that DNA repair genes are over-expressed in tumors from patients who do not respond to chemotherapy. This study highlights the importance of DNA repair gene expression in the progression of melanoma and gives insight into the biological basis of chemoresistance. It also provides preliminary evidence that this may be of predictive value in terms of resistance to chemotherapy. This article furthermore confirms that gene expression studies are possible on a large-scale using formalin-fixed tissue and may be used to identify further prognostic and predictive markers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Recruitment was facilitated by the UK National Cancer Research Network. We are grateful to May Chan, Clarisa Nolan, Susan Leake, Birute Karpavicius, Tricia Mack, Paul King, Sue Haynes, Elaine Fitzgibbon, Kate Gamble, Saila Waseem, Sandra Tovey, Christy Walker and Paul Affleck who collected and managed data.

This work was supported by Cancer Research UK (project grants C8216/A6129 and C8216/A8168), and program grants C588/A4994 and C588/A10589), and by the NIH (R01 CA83115). RJ is in receipt of a Bramall Fellowship and a Medical Research Council Clinical Research Training Fellowship (G0802123). KAP is funded by the Skin Cancer Research Fund. AM was funded by a grant from the Leeds Teaching Hospitals Trust Charitable Fund.

References

1. Balch CM, Buzaid AC, Soong SJ, et al. Final version of the american joint committee on cancer staging system for cutaneous melanoma. *J Clin Oncol*. 2001; 19:3635–48. [PubMed: 11504745]
2. Elder, D.; Murphy, G. Atlas of Tumor Pathology: Melanocytic Tumors of the Skin 2. Armed Forces Institute of Pathology; Washington: 1991. Malignant tumors (melanomas and related lesions); p. 103-205.third series
3. Cochran AJ, Elashoff D, Morton DL, Elashoff R. Individualized prognosis for melanoma patients. *Hum Pathol*. 2000; 31:327–31. [PubMed: 10746675]
4. Winnepeninckx V, Lazar V, Michiels S, et al. Gene expression profiling of primary cutaneous melanoma and clinical outcome. *J Natl Cancer Inst*. 2006; 98:472–82. [PubMed: 16595783]
5. Mandruzzato S, Callegaro A, Turcatel G, et al. A gene expression signature associated with survival in metastatic melanoma. *J Transl Med*. 2006; 4:50. [PubMed: 17129373]
6. Riker AI, Enkemann SA, Fodstad O, et al. The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis. *BMC Med Genomics*. 2008; 1:13. [PubMed: 18442402]
7. Kauffmann A, Rosselli F, Lazar V, et al. High expression of DNA repair pathways is associated with metastasis in melanoma patients. *Oncogene*. 2008; 27:565–73. [PubMed: 17891185]
8. Conway C, Mitra A, Jewell R, et al. Gene expression profiling of paraffin-embedded primary melanoma using the DASL assay identifies increased osteopontin expression as predictive of reduced relapse-free survival. *Clin Cancer Res*. 2009; 15:6939–46. [PubMed: 19887478]
9. Newton-Bishop JA, Beswick S, Randerson-Moor J, et al. Serum 25-hydroxyvitamin D3 levels are associated with breslow thickness at presentation and survival from melanoma. *J Clin Oncol*. 2009; 27:5439–44. [PubMed: 19770375]
10. Bibikova M, Talantov D, Chudin E, et al. Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays. *Am J Pathol*. 2004; 165:1799–807. [PubMed: 15509548]
11. Workman C, Jensen LJ, Jarmer H, et al. A new non-linear normalization method for reducing variability in DNA microarray experiments. *Genome Biol*. 2002; 3 research0048.
12. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ*. 1995; 310:170. [PubMed: 7833759]
13. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 2001; 25:402–8. [PubMed: 11846609]
14. Glaysher S, Yiannakis D, Gabriel FG, et al. Resistance gene expression determines the in vitro chemosensitivity of non-small cell lung cancer (NSCLC). *BMC Cancer*. 2009; 9:300. [PubMed: 19712441]
15. Helleday T, Petermann E, Lundin C, Hodgson B, Sharma RA. DNA repair pathways as targets for cancer therapy. *Nat Rev Cancer*. 2008; 8:193–204. [PubMed: 18256616]
16. Meric-Bernstam F, Gonzalez-Angulo AM. Targeting the mTOR signaling network for cancer therapy. *J Clin Oncol*. 2009; 27:2278–87. [PubMed: 19332717]
17. Agarwala S. Metastatic melanoma: an AJCC review. *Community Oncol*. 2008; 5:441–45.
18. Sarasin A, Kauffmann A. Overexpression of DNA repair genes is associated with metastasis: a new hypothesis. *Mutat Res*. 2008; 659:49–55. [PubMed: 18308619]

19. Schild D, Wiese C. Overexpression of RAD51 suppresses recombination defects: a possible mechanism to reverse genomic instability. *Nucleic Acids Res.* 2009
20. West SC. Molecular views of recombination proteins and their control. *Nat Rev Mol Cell Biol.* 2003; 4:435–45. [PubMed: 12778123]
21. Mortensen UH, Bendixen C, Sunjevaric I, Rothstein R. DNA strand annealing is promoted by the yeast Rad52 protein. *Proc Natl Acad Sci U S A.* 1996; 93:10729–34. [PubMed: 8855248]
22. Modrich P, Lahue R. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu Rev Biochem.* 1996; 65:101–33. [PubMed: 8811176]
23. Nitiss JL. DNA topoisomerase II and its growing repertoire of biological functions. *Nat Rev Cancer.* 2009; 9:327–37. [PubMed: 19377505]
24. Kosari F, Munz JM, Savci-Heijink CD, et al. Identification of prognostic biomarkers for prostate cancer. *Clin Cancer Res.* 2008; 14:1734–43. [PubMed: 18347174]
25. Wong N, Yeo W, Wong WL, et al. TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance. *Int J Cancer.* 2009; 124:644–52. [PubMed: 19003983]
26. Coss A, Tosetto M, Fox EJ, et al. Increased topoisomerase IIalpha expression in colorectal cancer is associated with advanced disease and chemotherapeutic resistance via inhibition of apoptosis. *Cancer Lett.* 2009; 276:228–38. [PubMed: 19111388]
27. Brase JC, Schmidt M, Fischbach T, et al. ERBB2 and TOP2A in breast cancer: a comprehensive analysis of gene amplification, RNA levels, and protein expression and their influence on prognosis and prediction. *Clin Cancer Res.* 2010; 16:2391–401. [PubMed: 20371687]
28. Glynn RW, Miller N, Kerin MJ. 17q12-21 - the pursuit of targeted therapy in breast cancer. *Cancer Treat Rev.* 36:224–9. [PubMed: 20100636]
29. Gogas HJ, Kirkwood JM, Sondak VK. Chemotherapy for metastatic melanoma: time for a change? *Cancer.* 2007; 109:455–64. [PubMed: 17200963]
30. Verbeek B, Southgate TD, Gilham DE, Margison GP. O6-Methylguanine-DNA methyltransferase inactivation and chemotherapy. *Br Med Bull.* 2008; 85:17–33. [PubMed: 18245773]
31. Hegi ME, Diserens AC, Godard S, et al. Clinical trial substantiates the predictive value of O-6-methylguanine-DNA methyltransferase promoter methylation in glioblastoma patients treated with temozolomide. *Clin Cancer Res.* 2004; 10:1871–4. [PubMed: 15041700]
32. Augustine CK, Yoo JS, Potti A, et al. Genomic and molecular profiling predicts response to temozolomide in melanoma. *Clin Cancer Res.* 2009; 15:502–10. [PubMed: 19147755]

Figure 1. Kaplan Meier curves of relapse-free survival for *RAD52* (A) and *TOP2A* (B) gene expression

Survivor functions have been estimated for each quartile of gene expression. Hazard ratios compared with the lowest quartile increase for each successive quartile. Analysis is adjusted for study and SNB status.

Table 1
Top 25 genes with expression levels in tumors most strongly related to RFS (adjusted for study and SNB status)

DNA repair genes or those associated with this process are highlighted in dark grey and genes involved in the cell cycle or cell proliferation are highlighted in light grey. Hazard ratios for reduced RFS are calculated for doubling of gene expression value. P-values are from the proportional hazards model.

Gene	Mean fold change between relapsers and non-relapsers	Hazard ratio	95% confidence interval	Significance value
<i>RAD51</i>	1.22	2.98	1.84-4.83	8.80×10^{-6}
<i>TK1</i>	1.12	4.68	2.29-9.58	0.00002
<i>ING1</i>	1.11	5.87	2.59-13.32	0.00002
<i>RAD52</i>	1.16	4.49	2.24-9.02	0.00002
<i>TFAP2C</i>	0.83	0.49	0.35-0.69	0.00004
<i>CCNA2</i>	1.16	2.66	1.67-4.23	0.00004
<i>BIRC5</i>	1.19	2.74	1.67-4.50	0.00007
<i>TOP2A</i>	1.12	3.84	1.96-7.52	0.00009
<i>CDH13</i>	0.80	0.48	0.33-0.71	0.0002
<i>HLF</i>	0.78	0.45	0.30-0.69	0.0002
<i>SPP1</i>	1.41	1.67	1.26-2.21	0.0003
<i>ITGB4</i>	0.79	0.58	0.43-0.79	0.0004
<i>FLI1</i>	0.83	0.51	0.35-0.75	0.0005
<i>MLF1</i>	1.15	2.74	1.55-4.83	0.0005
<i>EPHB4</i>	0.89	0.41	0.25-0.69	0.0008
<i>VEGFB</i>	1.04	9.84	2.57-37.71	0.0009
<i>ELK3</i>	0.95	0.20	0.07-0.53	0.001
<i>RAD54B</i>	1.18	2.08	1.33-3.25	0.001
<i>E2F1</i>	1.10	2.78	1.47-5.25	0.002
<i>WNT2</i>	1.34	2.49	1.41-4.41	0.002
<i>TERT</i>	1.29	1.79	1.24-2.58	0.002
<i>CDC2</i>	1.22	1.84	1.26-2.70	0.002
<i>CCNH</i>	1.08	3.39	1.58-7.27	0.002
<i>IFNGR1</i>	0.89	0.47	0.29-0.76	0.002
<i>PML</i>	0.96	0.21	0.07-0.58	0.003

Table 2
The association of genes involved in DNA repair with RFS in univariable and multivariable models

Part (i) of the table shows the association between single genes and survival. Data adjusted for study and SNB status only are presented in column 1. The association was adjusted for sex, patient age and tumor site (as known predictors of outcome) in column 2. In column 3, further adjustment is made for known histological predictors of outcome: Breslow thickness, mitotic rate and ulceration. Part (ii) of the table presents results of a multivariable Cox proportional hazards model with DNA repair genes identified as being predictive of RFS, Breslow thickness or mitotic rate (listed in Table 3). Significance values for genes that are independent predictors of survival are highlighted in bold. It can be seen that up-regulation of all three genes was associated with reduced RFS when adjusted for other known prognostic factors. *RAD52* and *TOP2A* remained significant in multivariable analysis.

Gene	(i) Association between single genes and relapse free survival						(ii) Multivariable model with all DNA repair genes predictive of RFS, Breslow thickness or mitotic rate					
	1. Analysis adjusted for study and SNB status only			2. Analysis further adjusted for age and sex of patient and site of tumor			3. Analysis further adjusted for Breslow thickness, ulceration and mitotic rate of tumor			1. Analysis adjusted for study and SNB status only		
	Fold	HR (95% CI)	P value	HR (95% CI)	P value	HR (95% CI)	HR (95% CI)	P value	HR (95% CI)	P value	HR (95% CI)	P value
<i>RAD51</i>	1.22	2.98 (1.84-4.83)	8.80×10^{-6}	3.22 (1.95-5.32)	4.67×10^{-6}	2.55 (1.46-4.47)	2.55 (1.46-4.47)	0.001	1.91 (0.98-3.72)	0.06	2.13 (1.09-4.18)	0.03
<i>RAD52</i>	1.16	4.49 (2.24-9.02)	0.00002	4.32 (2.12-8.82)	0.00006	2.79 (1.27-6.15)	2.79 (1.27-6.15)	0.01	4.73 (1.99-11.26)	0.0004	4.20 (1.77-9.95)	0.001
<i>TOP2A</i>	1.12	3.84 (1.96-7.52)	0.00009	3.87 (1.97-7.61)	0.00009	3.39 (1.62-7.11)	3.39 (1.62-7.11)	0.001	3.06 (1.33-7.06)	0.009	2.78 (1.20-6.45)	0.02
											3.15 (1.27-7.76)	0.01
											3.14 (1.23-8.00)	0.02

Table 3
DNA repair genes differentially expressed in relation to Breslow thickness and mitotic rate

Genes listed were identified as significantly differentially expressed ($p < 0.0001$) in tumors with greater Breslow thickness or mitotic rate using linear regression. Fold changes in gene expression are presented for each tumor thickness and mitotic rate group compared to baseline groups. Analyses are adjusted for study only and further adjusted for age of patient at diagnosis, site of tumor and sex of patient.

Gene	Breslow thickness			Mitotic rate		
	Fold changes in gene expression between indicated group and tumors <2mm thick		Significance level (adjusted for age and sex of patient and site of tumor)	Fold changes in gene expression between indicated group and tumors with mitotic rate <1/mm ²		Significance level (adjusted for age and sex of patient and site of tumor)
	2-4 mm	>4 mm		1-6/mm ²	>6/mm ²	
<i>RAD54B</i>	1.20	1.39	1.24×10^{-8}	1.29	1.62	6.28×10^{-7}
<i>MSH2</i>	1.20	1.25	1.27×10^{-8}	1.22	1.37	0.00001
<i>RAD51</i>	1.18	1.33	9.18×10^{-8}	1.22	1.53	7.65×10^{-8}
<i>RAD52</i>	1.07	1.21	9.18×10^{-8}	1.07	1.16	0.009
<i>CHEK1</i>	1.21	1.36	4.63×10^{-6}	1.13	1.42	0.0004
<i>TOP2A</i>	1.07	1.19	0.00003	1.21	1.38	8.87×10^{-9}
<i>RAD54L</i>	1.20	1.29	0.00004	1.09	1.31	0.004
<i>BRCA1</i>	1.20	1.25	0.00005	1.18	1.48	0.00001
<i>BRCA2</i>	1.18	1.20	0.00009	1.22	1.32	0.004
<i>MSH6</i>	1.04	1.08	0.0008	1.07	1.15	1.28×10^{-6}

Table 4
Genes most correlated with principal component 1

Correlation coefficients adjusted for study. These 37 genes with a correlation coefficient significant at $p < 10^{-10}$ were used in IPA to infer cellular and molecular functions and to build gene networks.

Gene	Correlation	P value
<i>CDC2</i>	0.54	1.3×10^{-27}
<i>TYMS</i>	0.51	8.4×10^{-25}
<i>BIRC5</i>	0.50	1.2×10^{-23}
<i>HMMR</i>	0.49	1.8×10^{-22}
<i>CCNA2</i>	0.48	9.0×10^{-22}
<i>CDC25A</i>	0.48	1.1×10^{-21}
<i>TK1</i>	0.48	1.7×10^{-21}
<i>PCNA</i>	0.45	2.8×10^{-19}
<i>CDC25C</i>	0.45	5.6×10^{-19}
<i>DEK</i>	0.45	8.6×10^{-19}
<i>DCN</i>	-0.44	5.0×10^{-18}
<i>CDK4</i>	0.42	7.1×10^{-17}
<i>WEE1</i>	0.42	9.9×10^{-17}
<i>ETS2</i>	-0.41	4.9×10^{-16}
<i>PDGFRA</i>	-0.41	8.9×10^{-16}
<i>BLM</i>	0.40	7.2×10^{-15}
<i>ITGB4</i>	-0.39	2.4×10^{-14}
<i>CDKN2C</i>	0.39	2.4×10^{-14}
<i>E2F1</i>	0.38	1.5×10^{-13}
<i>XRCC2</i>	0.38	2.4×10^{-13}
<i>MYBL2</i>	0.37	3.8×10^{-13}
<i>FYN</i>	-0.37	5.7×10^{-13}
<i>XRCC4</i>	0.37	1.2×10^{-12}
<i>PDGFRB</i>	-0.37	1.3×10^{-12}
<i>BCL6</i>	-0.36	2.8×10^{-12}
<i>CDH11</i>	-0.36	4.0×10^{-12}
<i>FOSL2</i>	-0.36	4.4×10^{-12}
<i>MAF</i>	-0.36	5.5×10^{-12}
<i>MLF1</i>	0.36	5.9×10^{-12}
<i>COL18A1</i>	-0.36	6.3×10^{-12}
<i>PBX1</i>	-0.35	6.8×10^{-12}
<i>JUNB</i>	-0.35	1.1×10^{-11}
<i>IFNGR1</i>	-0.35	2.2×10^{-11}
<i>AXL</i>	-0.34	3.5×10^{-11}
<i>E2F3</i>	0.34	4.2×10^{-11}

Gene	Correlation	P value
<i>MCF2</i>	0.34	7.7×10^{-11}
<i>RECQL</i>	0.34	9.0×10^{-11}

Table 5
Genes differentially expressed in tumors unresponsive to chemotherapy using the CGEA-1 array

DNA repair genes on the array are listed followed by genes most over-expressed and under-expressed in unresponsive tumors. The number of samples from non-responders and responders is presented with fold change difference in gene expression between non-responders and responders.

	Gene	Number of non-responders/number of responders	Mean fold change in gene expression between non-responders and responders
DNA repair genes	<i>MSH6</i>	17/4	4.61
	<i>TOPO1</i>	21/4	3.15
	<i>MSH2</i>	16/4	3.06
	<i>XRCC1</i>	13/4	2.95
	<i>XRCC5</i>	22/5	2.87
	<i>TOP2A</i>	17/3	2.53
	<i>XRCC6</i>	22/5	2.47
	<i>MGMT</i>	9/2	2.21
	<i>RAD51</i>	17/3	2.13
	<i>ERCC1</i>	22/5	2.07
	<i>TOP2B</i>	22/4	1.96
	<i>XPA</i>	20/5	1.87
	<i>ATM</i>	18/3	1.81
	<i>BRCA1</i>	17/3	1.77
	<i>ERCC2</i>	18/4	1.73
	<i>GTF2H2</i>	21/4	1.51
	<i>MLH1</i>	16/2	1.21
Genes most over-expressed in non-responding tumors	<i>TAP4</i>	15/2	24.63
	<i>TS</i>	21/3	24.53
	<i>KI67</i>	9/2	12.06
	<i>MTII</i>	22/5	8.06
	<i>mTOR</i>	21/4	7.33
	<i>cN II</i>	19/5	6.62
	<i>PIK3CA</i>	21/4	5.02
	<i>MSH6</i>	17/4	4.61
	<i>HSP90</i>	22/5	4.51
	<i>SOD 1</i>	22/5	4.40
Genes most under-expressed in non-responding	<i>MCL1</i>	21/4	0.38
	<i>FAS</i>	16/4	0.45
	<i>FPGS</i>	12/2	0.71
	<i>BAX</i>	21/5	0.86
	<i>P53</i>	19/5	0.92

	Gene	Number of non-responders/number of responders	Mean fold change in gene expression between non-responders and responders
tumors			