Oncogene (2008) 27, 565–573

www.nature.com/onc

ORIGINAL ARTICLE

High expression of DNA repair pathways is associated with metastasis in melanoma patients

A Kauffmann[1], F Rosselli[1], V Lazar[2], V Winnepenninckx[3], A Mansuet-Lupo[4], P Dessen[1], JJ van den Oord[5], A Spatz[4] and A Sarasin[1]

[1]*Laboratory of Genomes and Cancer, FRE2939 CNRS, Gustave-Roussy Institute, Villejuif and University Paris-Sud, Orsay, France;* [2]*Unit of Functional Genomics, Gustave-Roussy Institute, Villejuif, France;* [3]*Department of Pathology, University Hospitals of Maastricht, P. Debyelaan 25, Maastricht, HX, The Netherlands;* [4]*Department of Bio-Pathology, Gustave-Roussy Institute, Villejuif, France and* [5]*Department of Morphology and Molecular Pathology, University Hospitals, Katholieke Universiteit Leuven (KUL), Leuven, Belgium*

We have identified a gene-profile signature for human primary malignant melanoma associated with metastasis to distant sites and poor prognosis. We analyse the differential gene expression by looking at whole biological pathways rather than individual genes. Among the most significant pathways associated with progression to metastasis, we found the DNA replication ($P = 10^{-14}$) and the DNA repair pathways ($P = 10^{-16}$). We concentrated our analysis on DNA repair and found that 48 genes of this category, among a list of 234 genes, are associated with metastatic progression. These genes belong essentially to the pathways allowing recovery of stalled replication forks due to spontaneous blockage or induced DNA lesions. Because almost all these differentially expressed repair genes were overexpressed in primary tumors with bad prognosis, we speculate that primary melanoma cells that will metastasize try to replicate in a fast and error-free mode. In contrast to the progression from melanocytes to primary melanoma, genetic stability appears to be necessary for a melanoma cell to give rise to distant metastasis. This overexpression of repair genes explains nicely the extraordinary resistance of metastatic melanoma to chemo- and radio-therapy. Our results may open a new avenue for the discovery of drugs active on human metastatic melanoma.
Oncogene (2008) **27,** 565–573; doi:10.1038/sj.onc.1210700; published online 24 September 2007

Keywords: melanoma; metastasis; DNA repair; microarray

Correspondence: Professor A Sarasin, Laboratory of Genomes and Cancer, FRE2939 CNRS, Institut Gustave Roussy, PR2, 39 rue Camille Desmoulins, Villejuif 94805, France.
E-mail: sarasin@igr.fr
Availability: http://www.ebi.ac.uk/arrayexpress
E-TABM-1 IGR_ MELANOMA_STUDY
E-TABM-2 IGR_MELANOMA_VALID
Received 29 March 2007; revised 12 June 2007; accepted 22 June 2007; published online 24 September 2007

Introduction

Cutaneous melanoma develops from the malignant transformation of melanocytes that represent 1–2% of the total cell population found in the human epidermis. Melanoma causes a considerable public health burden because of its dramatic rise in incidence of about 3–8% per year in Caucasian populations worldwide since the mid-1960s (Thompson *et al*., 2005). The melanoma incidence rose 5% per year between 1978 and 2000 in France and increased for the whole of the 1979–1998 period by 130% in the United States (Howe *et al*., 2001; Jemal *et al*., 2001). Up to now, no active targeted therapy has been successful. Indeed, most patients with metastatic disease at distant sites show a median survival of only 4–6 months (Balch *et al*., 2001).

Genetic predisposition to melanoma is very well characterized but represents only 5–10% of all reported cases (Chin *et al*., 2006). Sunlight exposure represents a major environmental risk factor as exemplified by the increase of melanoma incidence in relation to latitude for fair-skinned people. Similarly, light-skinned individuals with blond or red hair, who freckle or sunburn easily, exhibit a much higher risk of developing melanoma than individuals with darker skin. The best evidence of a causative link with sun exposure is the association of melanoma with severe episodic sunburn in infancy (Armstrong and Kricker, 2001). However, it is clear that melanoma can also develop in parts of the body not exposed to sunlight, and that other genetic and intrinsic factors play a role.

Exposure to the UV-spectrum of sunlight induces DNA damage that is mainly repaired by the nucleotide excision repair (NER) pathway avoiding the induction of mutation and cancer. A low level of DNA repair is associated with increased risk of skin cancer including melanoma. The NER-deficient xeroderma pigmentosum (XP) patients develop numerous skin cancers following UV-exposure including a 2000-fold higher incidence of melanoma as compared to DNA repair-proficient individuals (Spatz *et al*., 2001; Stary and Sarasin, 2002). In the absence of repair, DNA replication is stalled by UV-induced DNA lesions that can be eventually bypassed by

specialized DNA polymerases. Depending upon the translesion DNA polymerase used to replicate a given DNA lesion, this process could be mutagenic and eventually lead to cancer. For example, XP variants, who are deficient in the translesion polymerase η, will replicate the lesions with a more error-prone polymerase, will produce more mutations following sun exposure and will develop more skin cancers, including melanoma (Spatz *et al*., 2001).

To understand the progression of primary melanoma to distant metastasis, we used a large bank of frozen primary melanomas in order to correlate gene expression profiles with distant metastasis-free survival at 4 years (Winnepenninckx *et al*., 2006). We identified 254 genes that allowed us to differentiate between primary tumors that will metastasize within 4 years (M+) versus those that will not (M−) (Winnepenninckx *et al*., 2006). This kind of analysis is not only useful at the single gene level but is also extremely informative to characterize the pathways that are altered as a consequence of the neoplastic process as well as on the pathways that ensure tumor growth and invasiveness. Consequently, instead of looking at individual genes, we decided to look for whole biological pathways linked to metastasis. The pathways, mostly associated with metastatic risk appeared to be those involved in replication licensing and origin firing, in cell cycle control and DNA repair. Because sun exposure is widely regarded as the critical environmental risk factor for melanoma and because the efficacy of DNA repair of UV-induced DNA lesions is directly responsible for cancer prevention, we looked for expression of DNA repair genes and more generally genes involved in the maintenance of genetic stability in relation to the risk of developing distant metastasis or survival.

By applying a novel classification of genes by functional pathways, we could demonstrate that the genes involved in repair, and generally in governing the stability of the cellular genome, are highly overrepresented among the genes that are differentially expressed between primary melanomas that will metastasize over a 4-year-period and those that will not. The main genes correlated with progression to metastasis are those implicated in the recovery of replication fork stalling and recombination. This implies that the majority of neoplastic cells, found in primary melanomas that will metastasize, have overexpressed their genes responsible for faithful and efficient repair pathways, ultimately resulting in genetically-stable cells that are able to metastasize and grow at distant sites.

Results

To understand melanoma progression, we analysed the gene expression profiles of primary melanoma giving rise to metastasis (M+; $n=26$) versus the melanoma where no metastasis (M−; $n=34$) was documented after a 4-year follow-up period (Winnepenninckx *et al*., 2006) (Supplementary Table 1). By applying searching for a biological interpretation of microarray experiments (SBIME) to the entire data set with Gene Ontology (GO) annotations, we found that two biological processes, that is, replication and repair, are especially overrepresented among the genes that are significantly differentially expressed between M+ and M− (Supplementary Table 2). By analysing these results, we noticed that some of the genes identified do not really belong to repair or replication pathways. Surprisingly, some well-known repair and replication genes were not found in these databases. We thus developed a novel classification by manual curation of the lists of genes from the replication and repair GO categories. The resulting lists are provided in Supplementary Table 3. By adapting SBIME with these new lists of genes, we re-analysed the data and observed significant overrepresentation of the replication and repair categories with *P*-values of 4.02×10^{-14} and 1.4×10^{-16}, respectively (Table 1 and Figure 1).

As can be seen in Figure 1, the vast majority of genes in both the repair and replication pathways are overexpressed in M+ versus M− tumors. An interesting underexpressed gene is NFIA that obtained an ANOVA *P*-value of 1.1×10^{-2} (which is just higher than the selected threshold of 1.0×10^{-2}) and that is found in the replication pathway because it has been associated with EBV replication (Gronostajski, 2000). However, this gene is also considered as a transcription factor and has been found to show tumor suppressor activity (Schuur *et al*., 1995). When the expression of this gene is analysed according to the thickness of the melanomas (that is the most significant prognostic parameter in primary melanoma of the skin), its expression decreases proportionally to the tumor thickness, linking its role as tumor suppressor to anti-metastatic activity (Figure 2b). It shows, moreover, that our experimental approach allows us to detect gene underexpression as well as overexpression between M+ and M− tumors. Because the direct relationship between the risk of metastasis and the expression of some replication factors has already been discussed in a previous paper (Winnepenninckx *et al*., 2006), we are only discussing here the role of DNA repair genes for metastasis formation.

The 44 genes that are overexpressed in M+ versus M− primary tumors (Table 1) can be classified in five general repair pathways.

Among the NER pathway, four genes are overexpressed that belong to the TFIIH factor (*GTF2H2*, *GTF2H3*) and to the replication pathway (*RFC*, *RPA*) (Table 1). Since the TFIIH factor is also involved in RNA transcription initiation, we can speculate that the NER pathway is only slightly implicated in metastasis induction.

Two interesting genes belonging to the Base Excision Repair pathway (BER) are overexpressed in M+ primary tumors (Table 1). The *hOGG1* gene, which encodes the main glycosylase, involved in the repair of oxidized guanines and, particularly, 8-oxo-guanine induced in skin by free radicals and UVA-exposure. The *NEIL3* gene (Rosenquist *et al*., 2003) which has strong homologies with NEIL1 and NEIL2 that also

Table 1 Comparison between the number of differentially expressed genes in melanoma from patients who will relapse (M+) versus patients who will not (M−) by ANOVA ($P<0.01$) in a particular category of pathway and the number of genes that would be selected by chance

Category	*Ng*	*Ngc*	*Ns*	*%*	P	*Regulation*	*Genes*				
Repair	234	221	48	22	1.4×10^{-16}	M+ > M−	NEIL3	RFC4	PTTG1	MSH6	OGG1
							CHEK1	RAD54L	TOP2A	RRM2	CHEK2
							EXO1	SMC2L1	DCLRE1A	RAD51	PCNA
							SUMO1	RFC5	RPA3	FANCG	MAD2L1
							XRCC5	POLE4	BLM	MSH2	BRCA1
							TERF1	EME1	RAD51AP1	SMC4L1	XRCC2
							GTF2H3	CCNH	RAD18	BRIP1	RAD17
							FANCD2	GTF2H2	PMS2L3	POLQ	TDP1
							RFC2	FANCA	HUS1	CHAF1A	
						M− > M+	CRY2	TERF2	FLJ35220	N4BP2	
Replication and polymerase	148	143	35	24	4.02×10^{-14}	M+ > M−	GMNN	RFC4	PTTG1	MCM6	PTTG2
							CDC6	Pfs2	CENPF	RRM2	MCM4
							SMC2L1	CDC45L	MCM3	PCNA	RFC5
							RPA3	TFAM	CTBP2	POLE4	BLM
							MSH2	PPIA	BRCA1	SMC4L1	ORC4L
							DEK	RAD17	ORC6L	HBXIP	MCM7
							POLQ	RFC2	CHAF1A		
						M− > M+	TOM1L2	NFIA			

Abbreviations: ANOVA, analysis of variance; Ngc, the number of corresponding significant genes present on the chip; Ng, the number of genes indexed in each functional category by our novel classification system; Ns, the number of genes with a significant difference in expression level between the two tumor groups. The column '%' shows the percentage of significantly expressed genes (Ns/Ngc × 100). '*P*' is the Fisher's exact test *P*-value for the whole pathway. The column 'Regulation' indicates the overall sense of variation of gene expression for the two classes of tumors. The genes numbered in Ns are then listed in decreasing order of the *P*-value (left to right).

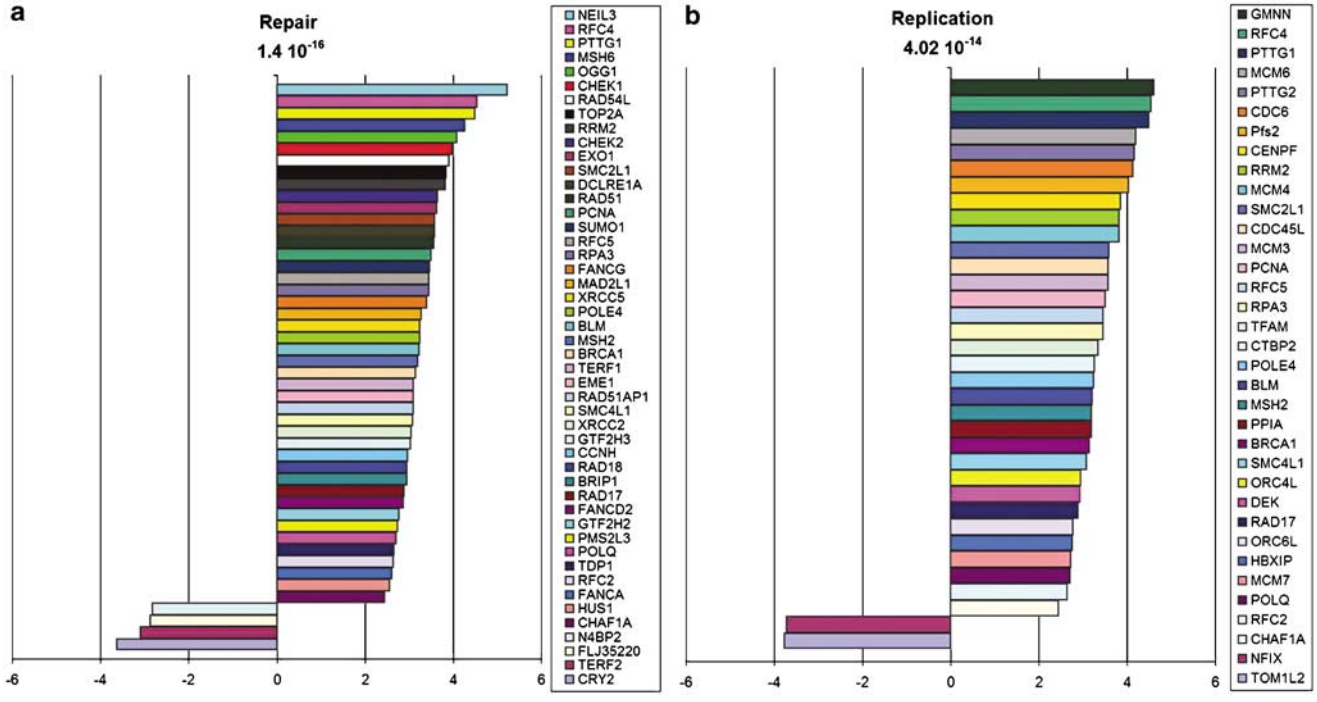

Figure 1 Standardized mean differences of log(ratio) between M− and M+ tumors for genes involved in DNA repair or DNA replication. A positive mean difference indicates overexpression in M+ compared to M−.

code for DNA glycosylases able to repair oxidized purines. Although the exact role of NEIL3 is not yet known, it is clear that overexpression of proteins that are able to recognize and repair oxidized guanines is associated with metastatic potential of melanoma. These oxidized bases are easily formed in sun-exposed and free-radicals exposed skin cells as well as in rapidly cycling cells.

As shown in Table 1, several mismatch repair proteins are overexpressed in M+ melanoma. MSH2 and MSH6

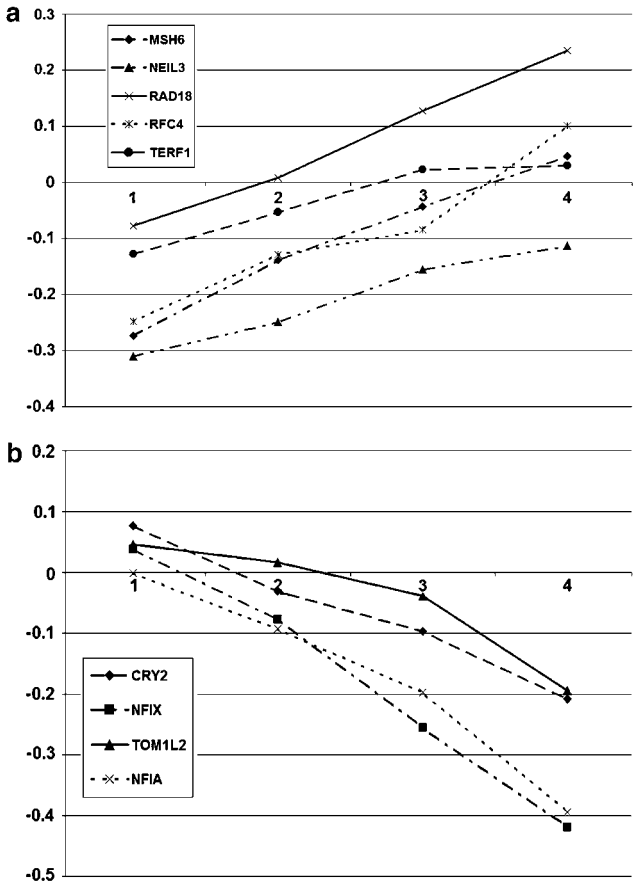

Figure 2 Mean log(ratio) in function of the four classes of the Breslow Index according to data given in the Supplementary Table 1. (**a**) Example of five repair or replication genes, which expression increases with the Breslow Index. (**b**) Example of four repair or replication genes, which expression decreases with the Breslow Index.

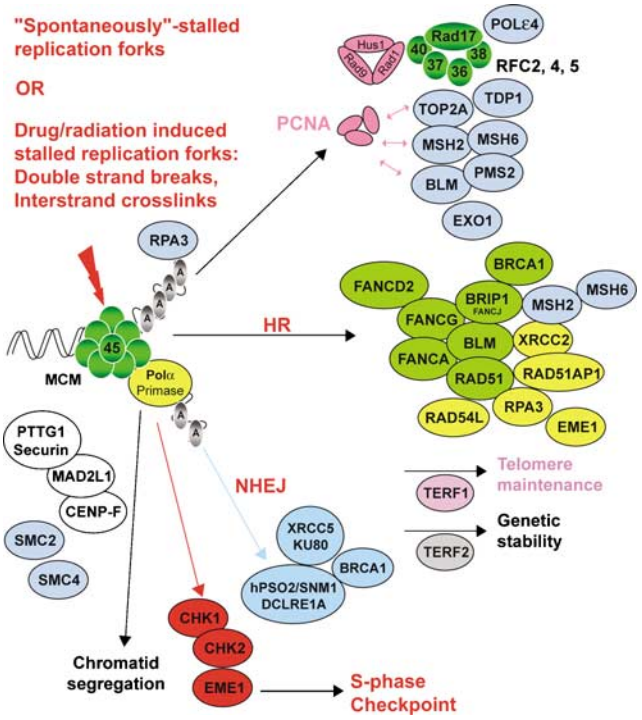

Figure 3 General scheme of the DNA repair genes overexpressed in M+ versus M− tumors and belonging to the surveillance of replication forks. All the listed genes are overexpressed in M+, except *TERF2*. It is believed that these genes are necessary to overcome the stalling of replication forks for both spontaneous-induced and drug-induced events.

proteins are able to recognize DNA mismatches produced during normal DNA replication but probably also to detect bases mispaired with unrepaired DNA lesions to avoid mutagenesis through translesion synthesis (TLS). Several genes overexpressed in M+ tumors are involved in (TLS) (Table 1), such as *RAD18*, *UBE2N* and *SUMO1*. The precise role of TLS is not clear, but according to the type of TLS polymerases and the type of DNA lesions, this pathway could be error-free or error-prone. The specialized pol Q protein belongs to the TLS class of DNA polymerases although we do not know yet its precise activity *in vivo*.

Among the 48 repair genes differentially expressed between M+ and M− primary tumors, 32 (67%) could be classified as belonging to pathways surveying replication forks to prevent or recover stalled replication forks (Figure 3, Table 1). Thirty-one out of the thirty-two are overexpressed. Blockage or formation of abnormal replication forks can occur spontaneously in cells exhibiting a high level of replication or following the induction of endogenous or exogenous DNA lesions. Double-strand breaks and inter-strand crosslinks are the major DNA lesions recognized and repaired by these pathways. Full repair of these stalled replication forks is necessary to ensure normal and error-free survival of rapidly replicating cells. These genes and corresponding repair pathways will be detailed in the discussion section.

The same analysis has been performed on the validation set of melanomas derived from another hospital (IGR) also described in Winnepenninckx *et al.* (2006) (Supplementary Table 1). The 8 M− primary tumors were compared with those of the 9 M+ primary tumors. The repair pathway and the replication pathway again showed significant *P*-values of 6×10^{-3} and 5.3×10^{-3}, respectively (Table 2).

Validation of DNA chip data has been already done with these series of tumors both at the reverse transcription–PCR level and at the protein expression level using tissue microarrays based on a different set of melanoma (Winnepenninckx *et al.*, 2006).

Discussion

By combining several databases as well as results published in the literature, we established a list of 234 genes involved in DNA repair (Supplementary Table 3). It is, however, probably more correct to consider this group in a broad sense, as constituted by genes whose products participate in the cell response to spontaneous

Table 2 Comparison between the number of differentially expressed genes in four different published studies concerning three different classes of primary tumors that will relapse (M+) versus primary tumors that will not (M−) by ANOVA ($P<0.01$) using the SBIME analysis (see legend of Table 1)

	Tumor types	*Repair*	*Replication*
Bladder (Dyrskjot *et al*., 2003) (27) $n=29$	Carcinoma	6.9×10^{-3}	NS
Breast (van't Veer *et al*., 2002) (28) $n=78$	Lymph-node negative primary tumor	1.4×10^{-2}	6.6×10^{-4}
Breast (Wang *et al*., 2005) (29) $n=286$	Lymph-node negative primary tumor	4.4×10^{-3}	8.0×10^{-3}
Squamous carcinoma of the oral cavity (O'Donnell *et al*., 2005) (30) $n=22$	Carcinoma	NS	5.1×10^{-4}
Melanoma (this study) $n=60$	Lymph-node negative primary tumor	1.4×10^{-16}	4.0×10^{-14}
Melanoma validation (this study) $n=17$	Lymph-node negative primary tumor	6.0×10^{-3}	5.3×10^{-3}

Abbreviations: ANOVA, analysis of variance; NS, nonsignificant; SBIME, searching for a biological interpretation of microarray experiments.

or induced DNA damage and are involved in the control of genetic stability *via* DNA repair and/or checkpoint regulation. Four major DNA repair pathways and four major cell cycle checkpoints have been described as: NER, BER, the MisMatch Repair (MMR) and recombinational repair (Homologous Recombination (HR) and Non Homologous End Joining (NHEJ)) and the G1/S, the intra-S-phase, the G2/M and the spindle checkpoints (Hoeijmakers, 2001). It is, however, clear that DNA repair and cell cycle checkpoint networks cooperate to ensure the maintenance of the genetic stability and the proper transmission of genetic information. In view of the tight connection between the two networks, several proteins participate in both, permitting a fine tuning of cell cycle progression allowing efficient DNA repair, resolution of the replication anomalies and correct separation and distribution of chromosomes in daughter cells.

Looking at the expression of genes coding for DNA repair/checkpoint proteins in M+ versus M− tumors, it appears clearly that no *bona fide* NER genes, such as those mutated in the inherited skin cancer prone syndrome XP, are differentially expressed. Overexpressed in M+ tumors were the genes coding for two proteins, GTF2H2 and GTF2H3 that both participate in the formation of TFIIH, a complex involved in NER and transcription. Overexpression of these two factors likely reflects an enhanced need for transcription as well as for DNA repair *per se* in M+ tumors.

The genes coding for BER proteins are also poorly represented. Only three elements of this repair pathway were found to be overexpressed in M+ tumors: NEIL3 that is still a putative DNA glycosylase, EXO1 that participates in BER as well as in MMR and recombination (see below) and OGG1, involved in the repair of oxidative DNA damage. Overexpression of *hOGG1*, the only *bona fide* BER gene, if followed by an increased enzymatic activity, could facilitate tumor growth and invasiveness, since base oxidation is the most frequent spontaneous and deleterious lesions observed in actively replicating cells.

In contrast to the rare occurrence of genuine NER and BER genes among the differentially expressed mRNA, factors involved in replication and post-replication events are statistically overrepresented among the genes overexpressed in M+ cells (Figure 1). The S-phase checkpoint and the post-replicative DNA repair mechanisms are essential for cell survival and proliferation in the presence of spontaneous and induced replicative stress. Cancer cells show genomic instability and a higher rate of DNA synthesis. Consequently, they can be considered to be in a continuous replicative stress situation. This condition correlates with the presence in cultured melanoma cells of low but spontaneous signs of DNA damage-induced or telomere-dependent response, such as histone H2AX phosphorylation (Halazonetis, 2004). This is the case for primary melanomas that show chromosomal instability (Bastian *et al*., 2003; Warters *et al*., 2005) and a higher frequency of γ-H2AX foci than non-transformed melanocytes (Warters *et al*., 2005). Consequently, keeping an elevated basal activity of both the intra-S-phase checkpoint and the S/G2-dependent DNA repair mechanisms could be important for survival and growth of metastatic melanoma cells. One of the few genes underexpressed in M+ tumors is *TERF2* that was originally described to be involved in the maintenance of telomere length. Two recent papers showed that overexpression of *TERF2* produces cellular hypersensitivity to UV and DNA crosslinking agents, as well as chromosome instability. In our study, *TERF2* is overexpressed in M− tumors leading to genetic instability and underexpressed in M+ tumors leading eventually to increased genetic stability (Bradshaw *et al*., 2005; Munoz *et al*., 2005).

One of the most differentially expressed genes between M+ and M− melanoma is *TOP2A* that belongs to the DNA repair/checkpoints group. It codes for an essential enzyme required during replication and chromosome segregation for relieving DNA torsional stress (Burden and Osheroff, 1998; Wang and Eastmond, 2002). TOP2A expression increases with cell proliferation as an effect rather than as a cause. However, Eder *et al*. (1995) demonstrated that cell lines ectopically overexpressing *TOP2A* are 5–10 times more resistant to alkylating agents; moreover, overexpression of *TOP2A* and its interacting protein PCNA was correlated with platinum-based therapy resistance in ovarian cancer. Interestingly, PCNA itself is overexpressed in M+ tumors as well as several other mRNA's coding for proteins that are able to interact with it. Among these are three components of the MMR system, that is, EXO1, MSH2 and MSH6. The

apparently coordinated selection of cells overexpressing mRNA's for PCNA and MMR components suggests that M+ cells could be armed to facilitate the repair of mis-incorporation errors generated during DNA replication. Since the inactivation of the MMR system leads to DNA damage hypersensitivity, it is likely that overexpression of the MMR system could improve the cellular resistance to DNA lesions. Similar to PCNA, other co-factors involved in DNA replication such as MCM-3, 4, 6 and 7, ORC4 and 6, RPA and RFC were found to be overexpressed in M+ tumors (Winnepenninckx *et al*., 2006). It is therefore likely that M+ cells exhibit a selective advantage from the overexpression of genes that allows a faithful DNA replication as well as better resistance to chemotherapy.

Globally, the majority of the differentially expressed genes emerging from our analysis code for proteins involved in rescuing stalled DNA replication forks, DNA double strand break (DSB)/interstrand crosslink (ICL) repair and telomere maintenance. Several M+ overexpressed mRNA's are components of the so-called FANC–BRCA pathway: FANCA, FANCG/XRCC9, FANCD2, FANCJ/BRIP1/BACH1, BRCA1, RAD51 and BLM. This pathway is mainly involved in ICL and DSB repair and in the rescue of blocked replication forks. Mutations in the FANC–BRCA pathway lead to both ionizing radiations (IR) and ICL hypersensitivity. The same is true for mutations in RAD54, DCLRE1A/SNM1, EME1 and XRCC2. The products of these genes, all overexpressed in M+ tumors, are not directly linked to the FANC–BRCA pathway but are clearly involved in ICL and recombinational repair. It has been reported that PCNA, RPA and MMR are also required for ICL removal, and the loss of their activities is associated with ICL hypersensitivity. A second subgroup of overexpressed mRNA in M+ melanomas includes the components of the clamp-loader RAD17-RFC2-5 (RAD17, RFC2, 4 and 5) as well as of the clamp RAD9-RAD1-HUS1 (HUS1). Both complexes are coordinately involved in DNA replication fork stall and rescue, recombination, NER and telomere maintenance (Niida and Nakanishi, 2006). A third subgroup of overexpressed genes in M+ melanomas includes the mRNA's coding for CHK1 and CHK2, two major protein kinases activated by ATM/ATR kinases to transduce signals induced by fork stalling, DNA DSB, DNA ICL or telomere abnormalities.

Mutations in all the genes included in the above reported subgroups are associated with (1) defects in transient DNA synthesis downregulation following DNA insults (DNA damage resistant DNA synthesis); (2) hypersensitivity to platinum compounds, IR and/or crosslinking agents; (3) human syndromes with cancer predisposition. Consequently, by opposition, it seems likely that their overexpression in tumors not only ensures better replication but also participates in the resistance of the cancer cells to DNA damage regimens. Finally, all of the isolated mRNA's present in the three subgroups code for proteins directly or indirectly involved in telomere metabolism and telomere length maintenance, a central event during tumor growth. Interestingly, TERF1, specifically involved in telomere maintenance (Smogorzewska *et al*., 2000), is also overexpressed in M+ tumor cells. The overrepresentation of genes involved in intra-S-phase and post-replicative events is in accordance with data showing high levels of histone H2AX phosphorylation in melanoma cells (Warters *et al*., 2005). The presence of γ-H2AX is generally considered as an indication of the presence of DNA DSB or of dysfunctional telomeres and both events seem to be involved in the higher frequency of γ-H2AX foci observed in cultured melanoma cells (Warters *et al*., 2005).

Finally, the genes of two proteins coordinately involved in the spindle checkpoint that assures the proper segregation of chromosomes during mitosis are overexpressed in M+ tumors. The first, PTTG1/securin encodes a protein primarily involved in the regulation of sister chromatid separation during cell division. The second gene, Mitotic Arrest Deficient 2 (MAD2), codes for a pivotal component of the spindle assembly checkpoint, and it regulates the destruction of securin. It has been reported that ectopic overexpression of PTTG1/securin in HEK293 cells results in increased cell proliferation and formation of tumors in subcutaneous injected mice (Hamid *et al*., 2005). The mitotic spindle assembly checkpoint includes the MAD1/MAD2 complex that is recruited by the centromere-associated protein F (CENP-F) that we have also classified among the replication pathway genes and is overexpressed in M+ compared to M− tumors (Table 1). CENP-F has recently been found to be involved in the recruitment of spindle assembly checkpoint proteins to the kinetochores of misaligned chromosomes (Costa, 2005). Depletion of CENP-F level leads to failure of chromosome alignment and chromosome segregation defect (Laoukili *et al*., 2005).

The molecular mechanisms underlying the overexpression of so many DNA repair genes are still unknown. These genes may either have a unique common regulator or the neoplastic cells have been selected according to independent overexpression of small groups of genes. It is possible that irresolvable DNA damage or aberrant replication structures, due to a very active DNA synthesis in M+ tumors, lead to specific activation of signaling pathways. The p53 pathway could be one of the actors since we found that almost 50% of overexpressed DNA repair genes are directly (16 out of 48 genes) or indirectly (18 out of 48 genes) interacting with p53 (Ingenuity, Mountain View, CA, USA, http://www.ingenuity.com/products/pathways_analysis.html).

The Breslow Index (BI) measures the thickness of the primary melanoma and is directly correlated with metastatic risk. The majority of DNA repair gene expression increases proportionally to this Index (Figure 2a), indicating that overexpression of most DNA repair genes is associated with metastasis development. However, for some DNA repair or DNA replication genes, the expression decreases with increasing tumor thickness (Figure 2b). It appears that only 55% of all differentially expressed genes are overexpressed in

Table 3 Fisher's exact test on the significance of the proportion of upregulated genes in the repair and replication pathways as compared to the total number of genes differentially expressed on the whole microarray

	Significantly expressed genes on the array		*Repair*		*Replication*	
Total	2297		48		35	
Up	1262	55%	44	92%	33	94%
Down	1035	45%	4	8%	2	6%
P-value			3.0×10^{-8}		3.7×10^{-7}	

M+ (Table 3) indicating that we are not looking at a non-specific overexpression of many proteins due to a higher proliferation rate, but rather, to a specific activation of targeted DNA repair pathways.

It seems likely that the overexpression of components of the FANC–BRCA pathway more specifically relies on the role of this pathway in the resolution of DNA-stalled forks, independently from the kind of DNA lesions encountered by the replication machinery. Nevertheless, tumors may take advantage of the overexpression of all these gene products to improve both their cellular response and survival following DNA damage as well as their capacity to invade internal organs. To become a primary melanoma, melanocytes should undergo some type of genetic instability leading to various gene amplifications, gene deletions and mutations (Chin *et al*., 2006). From our data, we can hypothesize that a primary melanoma cell that will metastasize, that is, be able to leave the tumor, enter the circulation, extravasate, survive and grow in a distant organ, needs to stabilize its genome in order to achieve this dangerous journey. Overexpression of recombination proteins may be one way to allow metastatic cells to survive.

By looking at data reported for bladder cancer (Dyrskjot *et al*., 2003), breast cancer (van't Veer *et al*., 2002; Wang *et al*., 2005) and squamous cell carcinoma of the oral cavity (O'Donnell *et al*., 2005) repair or replication pathways obtained significant *P*-values using SBIME analysis to compare primary tumors from patients who will develop metastases with those patients who will not develop metastases (Table 2). Analysis of these independent data allows us to conclude that overexpressions of DNA repair genes, and consequently, metastatic progression are not specific for melanoma but appears to reflect a general phenomenon in at least three other types of tumors.

Treatment with physical or chemical agents that target DNA is a useful approach to cure cancer. The rationale for this has been that normal (wt) cells in replicative activity are more sensitive to DNA damage than stationary G0/G1 cells, in spite of the apoptosis induced by wt-p53 induction. Loss of p53 pathway functionality in cancer cells increases the DNA-damage resistance during the G1 phase, suggesting that treatment directed to S-phase DNA could be more effective as anti-cancer therapy. Nuclear accumulation of wild-type p53 has been described in melanoma. Inhibition of p53 transcriptional activity occurs through PTTG1 interaction (Bernal *et al*., 2002). As shown in Figure 2b, PTTG1 is one of the most overexpressed genes in our melanoma series with bad prognosis.

In conclusion, our analysis demonstrates that melanomas with metastatic potential are associated to the overexpression of genes related to resolve anomalies linked to DNA replication, telomere maintenance and the presence of DNA lesions in a replicating genome. Our observations fit well with the extreme resistance seen in melanomas to chemo- and radio-therapy and we believe that this resistance is, at least partially, linked to the phenomena reported here. Treatments targeting intra-S-phase signaling players, such as PI3Ks or CHK1/2 proteins, combined with a DNA damage regimen, could be useful in metastatic melanoma therapy.

Material and methods

Selection of melanoma patients and gene expression profiling
Primary tumors from melanoma patients with available fresh frozen tissue were identified from the files of the pathology departments at University Hospitals in Leuven, Belgium and at the Gustave Roussy Institute in Villejuif, France. All patients were treated uniformly—that is, complete excision of the primary melanoma that was followed by re-excision with margins according to the thickness of the primary tumor—and none of the patients had undergone sentinel lymph node biopsy during the pathologic staging procedure. No other treatment has been given to these patients.

Sixty melanoma tumors from patients who developed a metastasis or not during the 4 years following the diagnosis of the primary tumors were included in this study. Seventeen tumors were used as validation set. Melanoma patients are described in Winnepenninckx *et al*. (2006). The clinical and histopathologic characteristics of the study and validation series are summarized in Supplementary Table 1. RNA extraction, microarray hybridization and expression value analysis were performed as described in Winnepenninckx *et al*. (2006), and expression value and experiments information are deposited on the Array Express data repository at the European Bioinformatics Institute (http://www.ebi.ac.uk/arrayexpress/) under the following accession numbers: E-TABM-1 IGR_ MELANOMA_STUDY and E-TABM-2 IGR_MELANOMA_VALID (validation set).

Classical analysis with SBIME
Firstly, SBIME performs ANOVA explaining the variation in logarithm gene expression (log(ratio)) by the distant metastasis-free survival status at 4 years after diagnosis of the patients (26 patients with metastasis, M+ versus 34 patients without, M−). Genes showing the most significant standardized difference of average expression in the M+ and M− groups will have the lowest *P*-value derived from ANOVA. Secondly, SBIME recovers annotation files from GO and eliminates GO categories, which do not have a corresponding gene on the chip. For each category, SBIME compares the proportion of genes with an ANOVA *P*-value lower than 1.0×10^{-2} within that category with what would be expected if there was no relationship between the category and our subset of genes (with the total proportion of significant genes on the

array, that is, genes with an ANOVA *P*-value lower than 1.0×10^{-2}) by a Fisher's exact test (see Supplementary Methods).

DNA repair genes analysed

We used several databases listing DNA repair genes in a broad sense, that is, governing the genomic stability and took the list published by Wood *et al*. (http://www.cgal.icnet.uk/DNA_Repair_Genes.html) as a working model. We derived 15 classes of DNA repair genes among 234 genes. The majority of them have known DNA repair activity, but some of them exhibit only sequence homology with DNA repair genes isolated in other species including bacteria or yeast. The list of these genes is found in Supplementary Table 2. The 148 replication genes were selected from GO (Ashburner *et al*., 2000) and BioCarta (http://www.biocarta.com) databases, and genes of known replicative and translesion DNA polymerases were added.

Acknowledgements

The authors thank the Institut Gustave Roussy, and particularly Professors Gilbert Lenoir and Thomas Tursz for their constant interest in this work and their financial support. Grants were obtained from the 'Cancéropole Ile-de–France' and the INCa (Paris, France). AK got a fellowship from the French Ministry of National Education and Research (University Paris VII Denis Diderot, France). We thank Stefan Michiels and Dr Leela Daya-Grosjean for their critical reading of the manuscript.

References

Armstrong BK, Kricker A. (2001). The epidemiology of UV induced skin cancer. *J Photochem Photobiol B* **63**: 8–18.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM *et al*. (2000). Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet* **25**: 25–29.

Balch CM, Soong SJ, Gershenwald JE, Thompson JF, Reintgen DS, Cascinelli N *et al*. (2001). Prognostic factors analysis of 17 600 melanoma patients: validation of the American joint committee on cancer melanoma staging system. *J Clin Oncol* **19**: 3622–3634.

Bastian BC, Olshen AB, LeBoit PE, Pinkel D. (2003). Classifying melanocytic tumors based on DNA copy number changes. *Am J Pathol* **163**: 1765–1770.

Bernal JA, Luna R, Espina A, Lazaro I, Ramos-Morales F, Romero F *et al*. (2002). Human securin interacts with p53 and modulates p53-mediated transcriptional activity and apoptosis. *Nat Genet* **32**: 306–311.

Bradshaw PS, Stavropoulos DJ, Meyn MS. (2005). Human telomeric protein TRF2 associates with genomic double-strand breaks as an early response to DNA damage. *Nat Genet* **37**: 193–197.

Burden DA, Osheroff N. (1998). Mechanism of action of eukaryotic topoisomerase II and drugs targeted to the enzyme. *Biochim Biophys Acta* **1400**: 139–154.

Chin L, Garraway LA, Fisher DE. (2006). Malignant melanoma: genetics and therapeutics in the genomic era. *Genes Dev* **20**: 2149–2182.

Costa RH. (2005). FoxM1 dances with mitosis. *Nat Cell Biol* **7**: 108–110.

Dyrskjot L, Thykjaer T, Kruhoffer M, Jensen JL, Marcussen N, Hamilton-Dutoit S *et al*. (2003). Identifying distinct classes of bladder carcinoma using microarrays. *Nat Genet* **33**: 90–96.

Eder JP, Chan VT, Ng SW, Rizvi NA, Zacharoulis S, Teicher BA *et al*. (1995). DNA topoisomerase II alpha expression is associated with alkylating agent resistance. *Cancer Res* **55**: 6109–6116.

Gronostajski RM. (2000). Roles of the NFI/CTF gene family in transcription and development. *Gene* **249**: 31–45.

Halazonetis TD. (2004). Constitutively active DNA damage checkpoint pathways as the driving force for the high frequency of p53 mutations in human cancer. *DNA Repair (Amst)* **3**: 1057–1062.

Hamid T, Malik MT, Kakar SS. (2005). Ectopic expression of PTTG1/securin promotes tumorigenesis in human embryonic kidney cells. *Mol Cancer* **4**: 1–13.

Hoeijmakers JH. (2001). Genome maintenance mechanisms for preventing cancer. *Nature* **411**: 366–374.

Howe HL, Wingo PA, Thun MJ, Ries LA, Rosenberg HM, Feigal EG *et al*. (2001). Annual report to the nation on the status of cancer (1973 through 1998), featuring cancers with recent increasing trends. *J Natl Cancer Inst* **93**: 824–842.

Jemal A, Devesa SS, Hartge P, Tucker MA. (2001). Recent trends in cutaneous melanoma incidence among whites in the United States. *J Natl Cancer Inst* **93**: 678–683.

Laoukili J, Kooistra MR, Bras A, Kauw J, Kerkhoven RM, Morrison A *et al*. (2005). FoxM1 is required for execution of the mitotic programme and chromosome stability. *Nat Cell Biol* **7**: 126–136.

Munoz P, Blanco R, Flores JM, Blasco MA. (2005). XPF nuclease-dependent telomere loss and increased DNA damage in mice overexpressing TRF2 result in premature aging and cancer. *Nat Genet* **37**: 1063–1071.

Niida H, Nakanishi M. (2006). DNA damage checkpoints in mammals. *Mutagenesis* **21**: 3–9.

O'Donnell RK, Kupferman M, Wei SJ, Singhal S, Weber R, O'Malley B *et al*. (2005). Gene expression signature predicts lymphatic metastasis in squamous cell carcinoma of the oral cavity. *Oncogene* **24**: 1244–1251.

Rosenquist TA, Zaika E, Fernandes AS, Zharkov DO, Miller H, Grollman AP. (2003). New insights for understanding the transcription-coupled repair pathway. *DNA Repair (Amst)* **2**: 581–591.

Schuur ER, Kruse U, Iacovoni JS, Vogt PK. (1995). Nuclear factor I interferes with transformation induced by nuclear oncogenes. *Cell Growth Differ* **6**: 219–227.

Smogorzewska A, van Steensel B, Bianchi A, Oelmann S, Schaefer MR, Schnapp G *et al*. (2000). Control of human telomere length by TRF1 and TRF2. *Mol Cell Biol* **20**: 1659–1668.

Spatz A, Giglia-Mari G, Benhamou S, Sarasin A. (2001). Association between DNA repair-deficiency and high level of p53 mutations in melanoma of Xeroderma pigmentosum. *Cancer Res* **61**: 2480–2486.

Stary A, Sarasin A. (2002). The genetics of the hereditary xeroderma pigmentosum syndrome. *Biochimie* **84**: 49–60.

Thompson JF, Scolyer RA, Kefford RF. (2005). Cutaneous melanoma. *Lancet* **365**: 687–701.

van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M *et al*. (2002). A gene-expression signature as a predictor of survival in breast cancer. *Nature* **415**: 530–536.

Wang L, Eastmond DA. (2002). Catalytic inhibitors of topoisomerase II are DNA-damaging agents: induction of chromosomal damage by merbarone and ICRF-187. *Environ Mol Mutagen* **39**: 348–356.

Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F *et al*. (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* **365**: 671–679.

Warters RL, Adamson PJ, Pond CD, Leachman SA. (2005). Melanoma cells express elevated levels of phosphorylated histone H2AX foci. *J Invest Dermatol* **124**: 807–817.

Winnepenninckx V, Lazar V, Michiels S, Dessen P, Stas M, Alonso SR *et al*. (2006). Gene expression profiling of primary cutaneous melanoma and clinical outcome. *J Natl Cancer Inst* **98**: 472–482.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)