Altered molecular pathways in melanocytic lesions

Altered molecular pathways identify different histotypes within melanoma progression

Maria Scatolini[1]*, Maurizia Mello Grand[1]*, Enrico Grosso[1], Tiziana Venesio[2], Alberto Pisacane[2], Antonella Balsamo[2], Roberta Sirovich[3], Mauro Risio[2] and Giovanna Chiorino[1].

1 - Lab of Cancer Pharmacogenomics, Fondo Edo Tempia, Biella (Italy)

2 - Unit of Pathology, Institute for Cancer Research and Treatment (IRCC), Candiolo, Torino (Italy)

3 - Department of Mathematics, University of Torino, 10123 Torino, (Italy)

Abstract

BACKGROUND: malignant cutaneous melanoma originates in melanocytes, the pigment-producing cells of the skin and eye. This entity is relatively rare compared to other skin cancers (< 5%), but still it is responsible for 80% of all skin cancer-related deaths.

AIMS: to identify molecular signatures of melanoma progression and to understand if dysplastic nevi can be or not considered as an intermediate step within such progression.

METHODS: excisional biopsies from 18 common melanocytic nevi (CMN), 8 primary radial growth phase melanomas (RGPM), 15 primary vertical growth phase melanomas (VGPM), 5 melanoma metastases (MTS) and 11 dysplastic nevi (DN) were collected and global gene expression profiling of the tissues was performed using whole genome oligo-microarrays. Differentially expressed genes for each progression step (CMN vs RGPM, RGPM vs VGPM and VGPM vs MTS) were identified, and validation of selected transcripts by qRT-PCR was performed on an independent cohort of fixed samples.

RESULTS and CONCLUSION: the comparison between CMN and RGPM shows an enrichment of Gene Ontology (GO) categories associated with inter and intra-cellular junctions, whereas the transition from RGPM to VGPM is characterized by the deregulation of WNT3, MAPK and AKT pathways. In this step, enrichment analysis underlies the alteration of biological processes linked to apoptosis. Upregulation of genes involved into DNA double-strand breaks repair and downregulation of genes principally related to cellular adhesion are observed in MTS. As far as DN, some genes controlling proliferation were found more expressed than in CMN. Overall, DN display an heterogeneous behaviour in respect to both RGPM and VGPM. In fact, relevant oncogenes, such as MYC and BCL6, were less expressed than in RGPM, while, a similar modulation pattern of VGPM was detected for a subset of gene families, such as mismatch repair. This suggests that DN is not an intermediate step within melanoma progression, but a separate entity with an independent risk of progression to melanoma.

Introduction

Malignant cutaneous melanoma originates in melanocytes, the pigment-producing cells of the skin and eye. This entity is relatively rare compared with other skin cancers (< 5%), but still is responsible for 80% of all skin cancer-related deaths. Early-stage melanoma is a curable disease if properly excised, but the life span of patients with advanced disease is relatively short and little progress has been made in the treatment of metastatic melanoma (Lomas J et al. Frontiers in Bioscience 2008). Moreover, there are no available markers to predict which early stage tumors have a high risk of progression (Singh M. et al. British Journal of Dermatology 2008).

The initiation and progression of melanoma are characterized by a sequence of cyto-architectural modifications and stromal (dermis and subcutis) invasion on which the Clark's method of microstaging is based (Clark WH et al.Cancer Res,1969). The driving force behind the initiation and progression of melanoma development is the acquisition of somatic mutations in key regulatory genes. Vertical growth phase (corresponding to Clark's level III-V), which involves the escape from keratinocyte-mediated growth control, is the most critical event along this progression (Dahl C et al. APMIS 2007). The detection of a melanoma progression signature is fundamental to improve diagnosis and therapy of this aggressive pathology.

Microarrays for gene expression analysis are the best method to discover altered molecular mechanisms, as they permit to detect the expression levels of thousands of genes in one experiment (Chiorino G et al. J Biol Regul Homeost Agents 2008).

In this study we analyzed the gene expression profiling in melanocytic tumors obtained from skin biopsies, classified as: common melanocytic nevi, dysplastic nevi, radial growth phase melanomas, vertical growth phase melanomas and melanoma metastases. Global gene expression profiling of the tissues was performed using whole genome oligo-microarrays (Agilent Technologies) with a dye-swap duplication scheme. To delineate molecular signatures associated with melanoma progression, differentially expressed genes for each progression step were identified. Expression change validations were carried out on independent cohorts of samples. Furthermore, specific analyses focused on dysplatic nevi were performed to delineate their expression patterns and look for any relationship with the gene or altered pathway profiling of all the other analyzed lesions.

Materials and methods

Tissues and isolation of total RNA

All samples and clinical data were collected with the IRCC ethical commitee' s approval and patients' informed consent.

A total of 57 excisional biopsies from common melanocytic nevi (n = 18), dysplastic nevi (n = 11), primary radial growth phase malignant melanomas (n = 8), primary vertical growth phase malignant melanomas (n = 15) and melanoma metastases (n = 5) were analyzed. After removal, all the specimens were submitted to macroscopical and dermatoscopical examination, during which the pathologist picked up a small and superficial fragment of tissue containing the pigmented lesion. Then the selected areas were embedded into RNALater (RNA Stabilization Reagent – Ambion Inc, Austin TX) for further molecular analysis. Lesions were classified according to WHO-2003 classification of skin tumours (2006 edition). Staging was performed according to Clark's levels of invasion (table I).

All dysplastic nevi were characterized by the presence of both architectural disorders and cytological atypia as indicated in the WHO-2003 criteria. Atypia was graded as Low or High only by evaluating the cytology of melanocytes. Low grade DN contained mainly a population of small melanocytes and few, scattered large and oval cells with a large vesicular nucleus and a prominent nucleolus. The large cells were prevalent in high grade DN, but the absence of a pagetoid (intraepidermal) spread of melanocytes and the preservation of the peculiar architecture were fundamental for the differential diagnosis with a melanoma.

TableI : list of hybridized samples and their histological diagnosis. Paired codes indicate lesions from the same patient. Crosses besides sample codes refer to death from melanoma.

Code	Diagnosis	Atypia grade	Sex	Age	Code	Diagnosis	Clark's level	Sex	Age
1	CMN		F	57	27	RGPM	I	F	32
2	CMN		M	63	28	RGPM	I	M	71
3	CMN		M	57	29	RGPM	II	M	46
4a	CMN		M	41	30	RGPM	II	F	74
4b	CMN		M	41	7b	RGPM	II	M	42
5	CMN		F	26	31	RGPM	II	M	70
6	CMN		F	73	32	RGPM	II	F	54
7a	CMN		M	42	33	RGPM	II	F	50
8a	CMN		M	32					
9a	CMN		M	51	34	VGPM	III	F	55
10	CMN		M	39	35	VGPM	III	F	41
11	CMN		F	33	36 †	VGPM	III	M	62
12	CMN		F	43	37	VGPM	III	F	45
13	CMN		F	32	17b	VGPM	III	F	52
14	CMN		M	36	38	VGPM	IV	M	41
15	CMN		M	58	39	VGPM	IV	M	77
16	CMN		F	38	9b	VGPM	IV	M	81
17a	CMN		F	52	40	VGPM	IV	M	51
					41	VGPM	IV	F	56
18	DN	low	F	30	42	VGPM	IV	M	77
19	DN	low	F	23	43	VGPM	IV	M	55
20	DN	low	M	56	44	VGPM	V	F	81
21	DN	low	F	27	45	VGPM	V	M	76
22	DN	low	M	81	46	VGPM	V	M	76
23	DN	high	M	43	**Site**				
8b	DN	high	M	32	47 †	MTS	node	M	52
24	DN	high	F	36	48 †	MTS	node	M	58
25a	DN	high	M	43	49	MTS	derma	F	81
25b	DN	high	M	43	50 †	MTS	derma	M	81
26	DN	high	M	51	51	MTS	derma	F	66

All the samples were homogenized (Ultra – TurraxT8, Ika – Werke, Staufen, Germany) with 600 ul of RNA lysis buffer plus 1% 2-mercaptoethanol, and total RNA was isolated by using RNeasy Mini Kit (Qiagen, Dusseldorf, Germany). TotRNA quality was checked by means of RNA 6000 pico chip assays (Agilent Technologies, Palo Alto, CA) run on the Agilent 2100 bioanalyzer.

Moreover a new group of formalin-fixed and paraffin embedded samples divided into : 5 common nevi, 5 RGPM, 5 VGPM and 5 metastases was collected to validate melanoma progression signature. Total RNA from these samples was isolated using RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE Tissues (Ambion Inc, Austin TX).

Evaluation of apoptosis in dysplastic nevi

Apoptosis was evaluated on standard hematoxylin and eosin slides obtained from formalin-fixed, paraffin-embedded skin biopsy. The abundance of intracellular melanin pigment allowed us only to consider apoptotic bodies (AB) (Kerr 1972) as a reliable indicator of cell death. The apoptotic index was calculated by dividing the total number of AB by the total number of dysplastic cells and multiplying by 100.

Microarray probe preparation, hybridization and scanning

RNA isolated from each fresh tissue and from the human universal reference (BD™ Human Universal Reference Total RNA, Clontech, Palo Alto, CA) was amplified by means of the Amino Allyl MessageAmp I aRNA Kit (Ambion) to obtain amino allyl antisense RNA (aaRNA) following the method developed by Eberwine and coworkers. Two rounds of amplification were performed to obtain the necessary quantity of aaRNA for labeling. Briefly: mRNA was reverse transcribed in cDNA single strand; after the second strand synthesis (in the second round of amplification), cDNA was in vitro transcribed in aaRNA including amino allyl modified nucleotides (aaUTP). Both dsDNA and aaRNA underwent a purification step using columns provided with the kit. Labeling was performed using NHS ester Cy3 or Cy5 dyes (Amersham Biosciences, Buckinghamshire UK) able to react with the modified RNA. mRNA quality was checked by means of RNA 6000 nano chip assays (Agilent Technologies). At least 5 ug of mRNA for each sample were labeled and purified with columns. Equal amounts (0.75 ug) of labeled specimens from sample and reference were put together, fragmented and hybridized to oligonucleotide glass arrays representing 41K human unique genes and transcripts (Human Whole Genome Oligo Microarray Array, Agilent Technologies). All steps were performed using the In Situ Hybridization kit-plus (Agilent Technologies) and following the 60-mer oligo microarray processing protocol (Agilent Technologies). Then, slides were washed with the SSPE wash procedure and scanned with the dual-laser microarray scanner Agilent G2505B. For each sample, a dye-swap replicate was performed.

Microarray data analysis

Images were analyzed using Feature Extraction software (Agilent Technologies) version 7.6. Output files containing feature and background intensities and the related statistical parameters for red and green signals were then loaded into the Resolver SE System

(Rosetta Biosoftware, Seattle, WA) together with the scan images and the Agilent Human Whole Genome pattern file (version updated on August 7, 2006). Data processing and normalization were performed using the Agilent Human Whole Genome platform-specific error model. Replicated expression profiles were combined to form ratio experiments where each gene is associated to an expression fold-change and a p-value to assess the statistical significance of its modulation in the sample compared to the reference. New ratio experiments were also built by applying the "ratio split with common reference" option and calculating ratios over the average of non tumor samples (CMN). A "weight by error" divisive clustering algorithm available within the Rosetta Resolver system (with cosine correlation as similarity metrics) was applied to such new ratios to compare their expression profiles. At the same time, data treatment with the Limma (linear models for microarray data) package available within Bioconductor (http://www.bioconductor.org/) was also performed applying normexp background calculation, lowess within-array and A-quantile between-array normalization methods.

Gene selection and annotation

To compare class expression patterns, modified error-weighted one-way ANOVA available within the Resolver System for multi-class comparisons and empirical Bayes statistics within the Limma package for two class comparisons were applied to ratio experiments. Transcripts with significant modulation (p-value less than 0.01 for ANOVA and B value greater than 2 for empirical Bayes statistic) were further considered. A B value greater than 2 means that the gene has more than 88% probability of being differentially expressed (as given by the formula exp(2)/(1+exp(2)), and roughly corresponds to a moderate t-statistic adjusted p-value less than 0.01.

Several databases were interrogated to find out and compare functional information on single genes and on groups of differentially expressed genes. In particular, DAVID (http://david.abcc.ncifcrf.gov/) was used to look for the more affected biological processes and pathways.

Real Time RT- PCR

Quantitative RT polymerase chain reaction (qPCR) was performed on the independent cohort of fixed samples to validate a subset of differently expressed transcripts identified by microarray analysis on fresh samples. First strand cDNA synthesis was performed using 200 ng of total RNA and High-Capacity cDNA Reverse Transcription Kit (Applied

Biosystems, Foster City CA), according to the manufacturer's instructions.

The Real Time RT-PCR reactions were performed using the Opticon2 (Biorad, Hercules, CA) PCR system with Power SYBR Green PCR Master Mix (Applied Biosystems, Foster City CA). QuantiTect primer assays (Qiagen, Dusseldorf, Germany) and custom primers were used to perform qPCR reactions (tabella S-1).

PCR was done using the following protocol: initial denaturation 95°C for 15 min, then 40 cycles at 95°C for 15 sec for denaturation, 60°C for 30 sec (20°C/sec transition) for annealing, 72°C for 30 sec (2°C/sec transition) for extension. A melting step was performed after PCR to determine product purity. Copy number was determined using the crossing point (Cp) value, which is automatically calculated using the Opticon Monitor 3.1 software (Biorad). Calibration (standard) curves for each gene were constructed using serial dilutions (5 points, 5-fold dilution, duplicate) of cDNA from a Universal Reference (BD™ Human Universal Reference Total RNA, Clontech, Montain View, CA, USA) or genomic DNA (Human Genomic DNA male Promega, Madison, WI, USA).

The abundance of specific transcripts relative was established using a relatively stable transcript (Actb, beta actin) for normalization of input cDNA.

Results

In order to identify genes implicated during the course of melanoma development and progression, the gene expression patterns of common nevi (CN), dysplastic nevi (DN), radial growth phase melanoma (RGPM), vertical growth phase melanoma (VGPM) and melanoma metastasis (MTS) were compared.

Global gene expression profiling was performed using whole genome glass arrays and comparing each sample (with a dye-swap duplication scheme) to a common Universal reference. To visually inspect the relationships among samples, unsupervised clustering analysis was applied to the ratios between each DN, RGPM , VGPM or MTS sample and the average profile of all the CMN included into the study, just selecting transcripts with p-value less than 0.01 in at least half of the samples considered (Fig 1). Such approach revealed that MTS cluster apart, together with some VGPM, and that DN share some patterns with CMN, RGPM and, in some cases, to VGPM, but they also show a peculiar and specific expression profile.

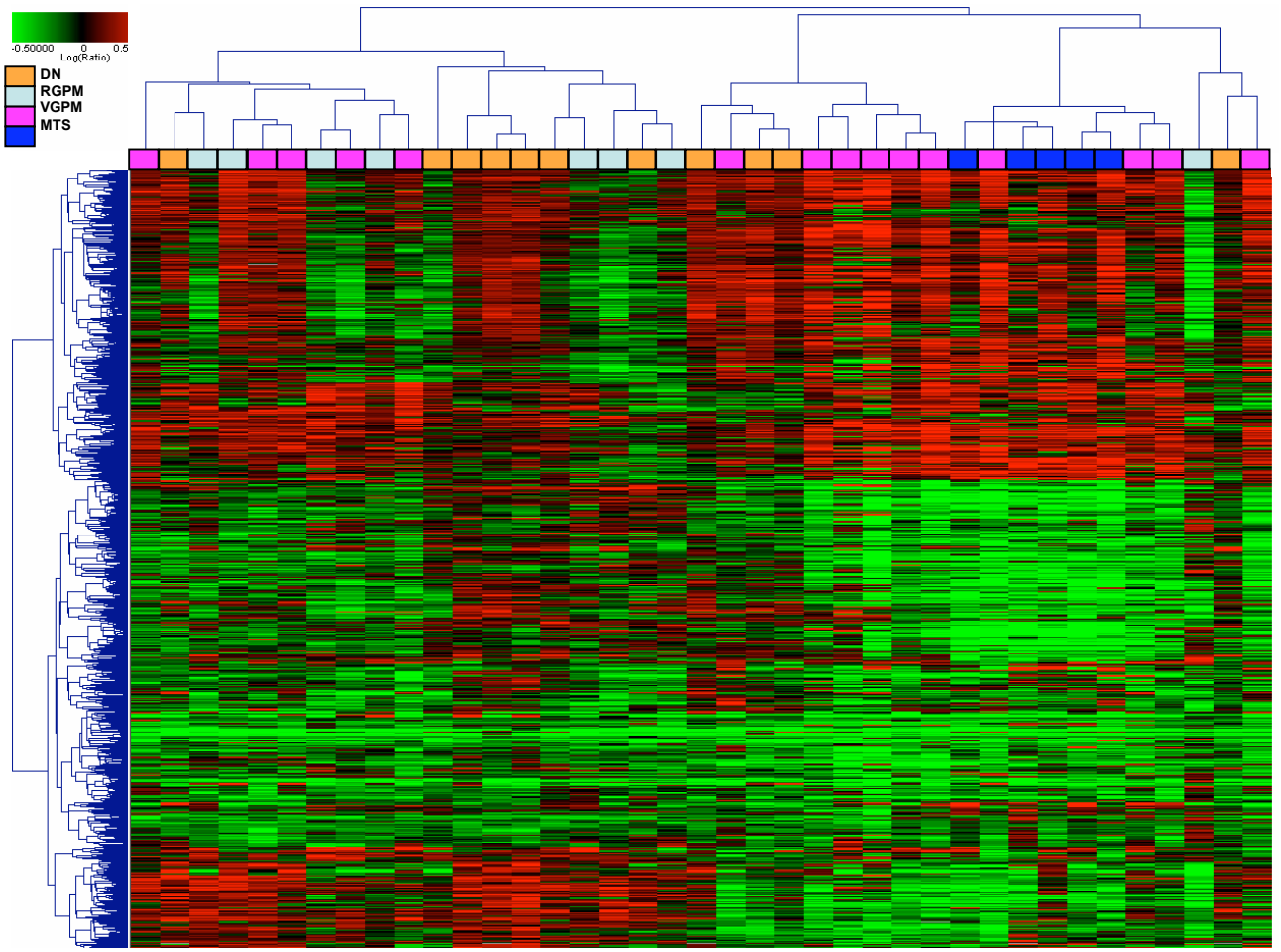

Figure1: " Weight by error" divisive clustering applied to the log10 expression ratios between each sample and the average of all CMN. Columns refer to 39 samples and rows to more than 3000 genes with a differential expression in at least 20 out of 39 samples (p-value less than 0.01).

Transcripts modulated in each progression step were then selected by applying a cut-off of 2 for the B value yielded by the limma procedure applied on original ratios (sample over common reference). Initially, DN were included into either the CMN class or the RGPM one, according to the degree of their cyto-architectural atypia (table I). This was done because there still exists a controversy on the identification of a distinct and specific category for such dysplastic lesions and also because unsupervised results did not show a clear class separation. However, the presence of such lesions among either CMN or RGPM made the distinction between these two classes and between RGPM and VGPM much less clear. Therefore, to prevent the increasing of heterogeneity, DN were excluded and this allowed the detection of many more differentially expressed transcripts between CMN and RGPM and, especially, between RGPM and VGPM. The great difference found between metastases and the rest of the samples was consistent with unsupervised results. Enrichment of Gene Ontology categories (in particular, biological processes) and gene networks or cellular pathways was performed on each list by using DAVID functional annotation tool and applying a cut-off of 0.01 on the enrichment score obtained.
For each progression step, some genes previously described as differentially expressed were found. Furthermore, other novel gene transcripts, never reported as involved in melanoma progression, were identified. The expression deregulation of some of these novel genes was then confirmed on an independent cohort of samples by RT-qPCR.

Common melanocytic nevus to radial growth phase (melanomagenesis)

To better characterize the melanomagenesis, differentially expressed genes between CMN and RGPM were analyzed. In this step only 36 annotated transcripts were found, as shown in table II. The biological processes of the gene ontology with a significant enrichment score during this step regard mainly cellular junction (intra and inter – cellular).
The expression changes of three transcripts were detected on an independent cohort of 10 samples (5 classified as CMN and 5 as RGPM), starting from archival fixed tissue (Fig 2).

Table II: transcripts differentially expressed between CMN and RGPM classes. Positive fold-change indicates up-regulation in RGPM and negative down-regulation. When two values are reported, the first one refers to microarray data (FC) while the second to RT-qPCR validation (RE = relative expression).

Gene symbol	RefSeq	Description	FC	RE
MGC12966	NM_001037163	Homo sapiens hypothetical protein MGC12966 (MGC12966), mRNA	-0.99	
HN1	NM_001002033	Hematological and neurological expressed 1	0.77	
NUDT11	NM_018159	Homo sapiens nudix (nucleoside diphosphate linked moiety X)-type motif 11 (NUDT11), mRNA	-0.92	
JARID2	NM_004973	Homo sapiens Jumonji, AT rich interactive domain 2 (JARID2), mRNA	0.55	
HSPC111	NM_016391	Homo sapiens hypothetical protein HSPC111 (HSPC111), mRNA	0.57	
NUDT10	NM_153183	Nudix -type motif 10	-1.39	-1.69
GLA	NM_000169	Homo sapiens galactosidase, alpha (GLA), mRNA	0.63	
HEY1	NM_012258	Hairy/enhancer-of-split related with YRPW motif 1	0.92	
LYPLA1	NM_006330	Lysophospholipase I	0.72	
GPM6A	NM_201592	Homo sapiens glycoprotein M6A (GPM6A), transcript variant 1, mRNA	-0.93	
GLIPR1	NM_006851	Homo sapiens GLI pathogenesis-related 1 (glioma) (GLIPR1), mRNA	0.88	
SLC35A1	NM_006416	Homo sapiens solute carrier family 35 (CMP-sialic acid transporter), member A1 (SLC35A1), mRNA	-0.64	
GDF15	NM_004864	Homo sapiens growth differentiation factor 15 (GDF15), mRNA	1.74	
HNRNPA1	NM_002136	Heterogeneous nuclear ribonucleoprotein A1	-0.93	
SCD5	NM_024906	Stearoyl-CoA desaturase 5	-1.13	
TRIM4	NM_033091	Homo sapiens tripartite motif-containing 4 (TRIM4), transcript variant alpha, mRNA	-0.76	
PTGDS	NM_000954	Homo sapiens prostaglandin D2 synthase 21kDa (brain) (PTGDS), mRNA	0.98	
RNASE1	NM_002933	Homo sapiens ribonuclease, RNase A family, 1 (pancreatic) (RNASE1), transcript variant 4, mRNA	1.06	0.48
ICAM5	NM_003259	Homo sapiens intercellular adhesion molecule 5, telencephalin (ICAM5), mRNA	0.75	
PKIA	NM_181839	Homo sapiens protein kinase (cAMP-dependent, catalytic) inhibitor alpha (PKIA), transcript variant 1, mRNA	-0.69	
LETMD1	NM_015416	LETM1 domain containing 1	-0.93	
RASA4	NM_006989	Homo sapiens RAS p21 protein activator 4 (RASA4), mRNA	0.73	
CADPS	NM_183394	Homo sapiens Ca2+-dependent secretion activator (CADPS), transcript variant 1, mRNA	0.7	
VAPA	NM_194434	Homo sapiens VAMP (vesicle-associated membrane protein)-associated protein A, 33kDa (VAPA), transcript variant 1, mRNA	-0.75	
CRISPLD2	NM_031476	Cysteine-rich secretory protein LCCL domain containing 2	0.65	
AMICA1	NM_153206	Adhesion molecule, interacts with CXADR antigen 1	1.14	1.06
SAP30L	NM_024632	SAP30-like	0.4	
ESAM	NM_138961	Homo sapiens endothelial cell adhesion molecule (ESAM), mRNA	0.67	
SORBS1	NM_001034954	Homo sapiens sorbin and SH3 domain containing 1 (SORBS1), mRNA	-	
CSRP2	NM_001321	Homo sapiens cysteine and glycine-rich protein 2 (CSRP2), mRNA	0.88	
C17orf45	NM_152350	Chromosome 17 open reading frame 45	-0.86	
CAMLG	NM_001745	Homo sapiens calcium modulating ligand (CAMLG), mRNA	-0.76	
WDR40A	NM_015397	WD repeat domain 40A	-0.37	
PRMT6	NM_018137	Homo sapiens protein arginine N-methyltransferase 6 (PRMT6), mRNA	-0.53	
RAB43	NM_198490	RAB43, member RAS oncogene family	0.53	
C14orf132	NM_020215	Chromosome 14 open reading frame132	-0.99	

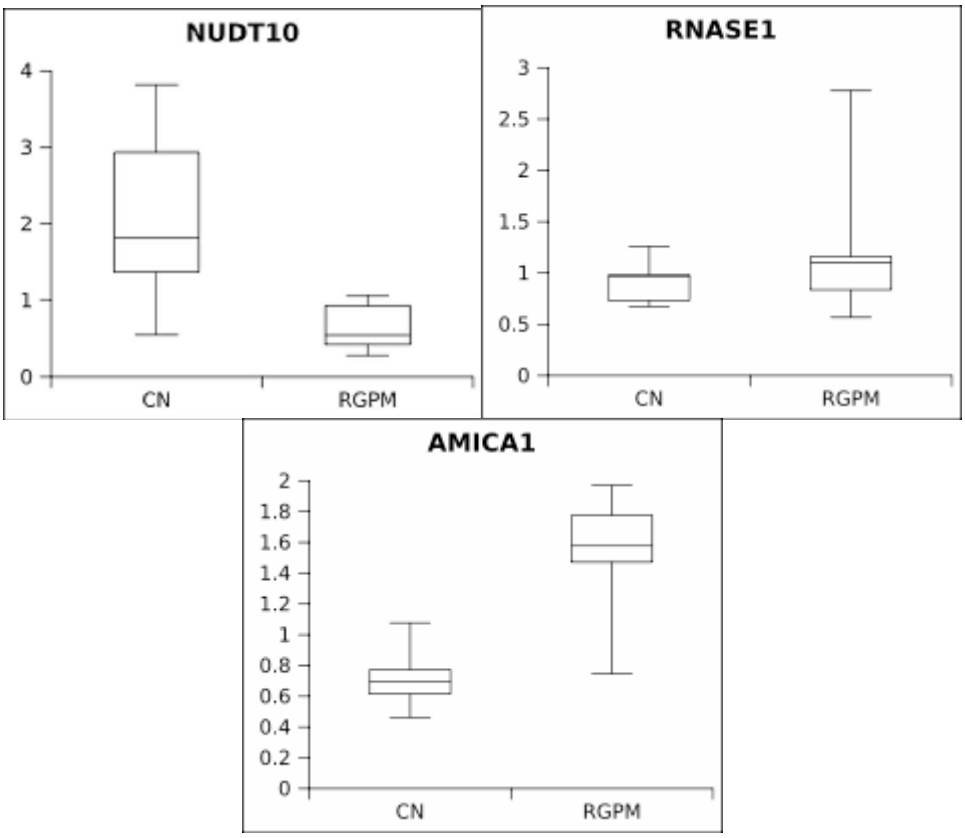

Figure 2: Levels of AMICA1, RNASE1 and NUDT10 mRNA in RGPM and CN samples. Box-plots represent the mRNA levels of (A) AMICA1, (B) RNASE1, (C) NUDT10 in the two series, normalized to ribosomal ACTB mRNA. Five samples were analyzed for AMICA1, RNASE1 and NUDT10 in each group. In the plots the first and third quantiles, the minimun, the median and maximum of relative expression mRNA levels are rapresented.

Radial to vertical growth phase

Much more sequences resulted differentially expressed in the transition from RGPM to VGPM (S-tableII). GO enrichment analysis separately applied on up and downregulated transcripts (table III) revealed that the main biological processes upregulated in RGPM are involved into steroid metabolism, regulation of transcription, induction of apoptosis, whereas in VGPM there is an upregulation of genes implicated into negative regulation of apoptosis and mismatch repair.

Table III: Biological processes level 5 of the Gene Ontology with significant enrichment score

Biological process level 5 term	genes	%	p-value
UP in RGPM			
steroid biosynthetic process	7	5,4	1,9 E - 5
sterol metabolic process	7	5,4	2,2 E -5
steroid metabolic process	9	7,0	2,7 E -5
sterol biosynthetic process	5	3,9	9,1 E -5
cholesterol metabolic process	6	4,7	1,6 E -4
positive regulation of transcription	8	6,2	7,4 E -3
positive regulation of cellular metabolic process	9	7,0	7,9 E -3
positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	8	6,2	8,6 E -3
positive regulation of transcription, DNA - dependent	7	5,4	9,9 E -3
lipid biosynthetic process	7	5,4	1,0 E -2
UP in VGPM			
negative regulation of apoptosis	11	3,8	4,3 E -3
negative regulation of programmed cell death	11	3,8	4,8 E -3
modification-dependent macromolecule catabolic process	9	3,1	6,1 E -3
mismatch repair	4	1,4	7,0 E -3
cellular protein catabolic process	9	3,1	7,0 E -3
maintenance of fidelity during DNA-dependent DNA replication	4	1,4	7,8 E -3

Quantitative RT-PCR validations were carried out on four of the differentially expressed transcripts, using the 5 fixed RGPM of the previously described independent assay, plus 5 fixed VGPM.

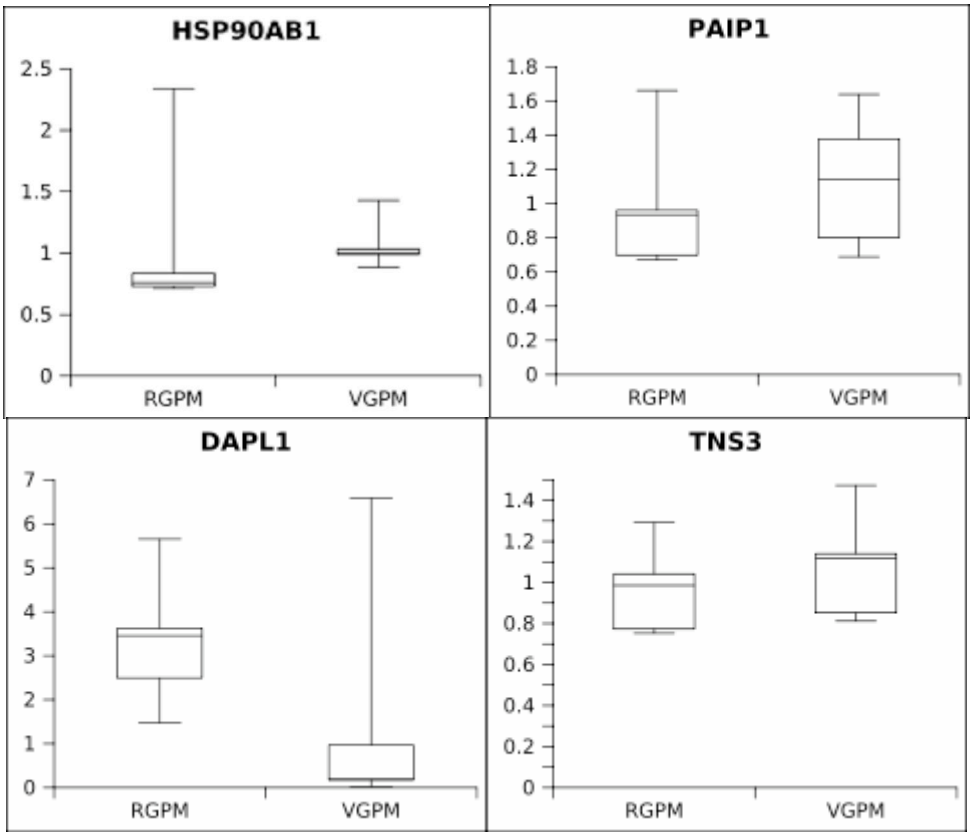

Figure 3: Levels of HSP90AB1, TNS3, PAIP1 and DAPL1 mRNA in VGPM and RGPM samples. Box-plots represent the mRNA levels in the two independent series (5 fixed RGPMs and 5 fixed VGPMs), normalized to ribosomal ACTB mRNA.

Vertical growth phase to metastasis

Despite the different locations at which the melanoma metastases analyzed in this study were taken, their gene expression profile resulted consistently and homogeneously altered with respect to VGPM, even if only 5 metastases were included into the analysis (S-tableIII). Reported in table IV are the biological processes mostly overrepresented within the list of differentially expressed genes of this progression step, divided according to the sign of the expression change. It is interesting to note that in VGPM there is an alteration of pathways implicated in development (ectoderm, epidermis, tissue) but also in cell differentiation (epidermal and epithelial). Differently, in MTS there is an alteration of pathways involved in cell cycle and DNA repair. Most of the deregulated transcripts are underexpressed in metastases and can be classified into a small number of families, as shown in table V. For this step, four genes were validated with RT-qPCR (Fig. 4)

Table IV: Biological processes level 5 of the Gene Ontology with significant enrichment score

Biological process level 5 term	genes	%	p-value
UP in VGPM			
ectoderm development	38	12,8	7,0 E -34
epidermis development	37	12,5	1,2 E -33
tissue development	41	13,8	1,9 E -24
epidermis morphogenesis	13	4,4	3,3 E -11
keratinization	11	3,7	6,3 E -11
epidermal cell differentiation	12	4,0	1,9 E -10
tissue morphogenesis	13	4,4	2,8 E - 9
epithelial cell differentiation	7	2,4	4,6 E -5
angiogenesis	8	2,7	8,2 E - 3
regulation of phosphatase metabolic process	6	2,0	8,5 E - 3
regulation of phosphorus metabolic process	6	2,0	8,5 E - 3
UP in MTS			
M phase	11	15,1	4,5 E -7
mitosis	9	12,3	7,2 E -6
M phase of mitotic cell cycle	9	12,3	7,7 E -6
regulation of progression through cell cycle	10	13,7	4,8 E - 4
cell cycle checkpoint	4	5,5	2,7 E - 3
DNA repair	6	8,2	5,3 E - 3
interphase	4	5,5	9,1 E - 3

Table V: Families of transcripts underexpressed in MTS vs VGPM

Family genes	Genes Symbol	Family genes	Genes Symbol
Laminins	LAMC2	**Cadherins**	CDH1
	LAMA3		CDH13
	LAMB3		CELSR1
Collagens	COL13A1	**Cystatine**	CSTA
	COL17A1		CST6
	COL4A6		CSTB
Kallikreins	KLK5	**Serpine**	SERPINEB3
	KLK7		SERPINE2B
	KLK8		SERPINEB5
	KLK10		SERPINEB13
	KLK11		SERPINEB7
Keratins	KRT6B	**Calcium binding proteins**	S100A8
	KRT10		S100A14
	KRT14		S100A7
	KRT23		S100A9
	KRT80		S100A10
	KRT17		S100P
	KRT33A		S100A2
	KRT6A	**Small proline rich proteins**	SPRR1B
	KRT5		SPRR1A
	KRT2A		SPRR2D
	KRT16		SPRR2E
	KRT15	**Desmocollins**	DSC3
	KRT6E		DSC1

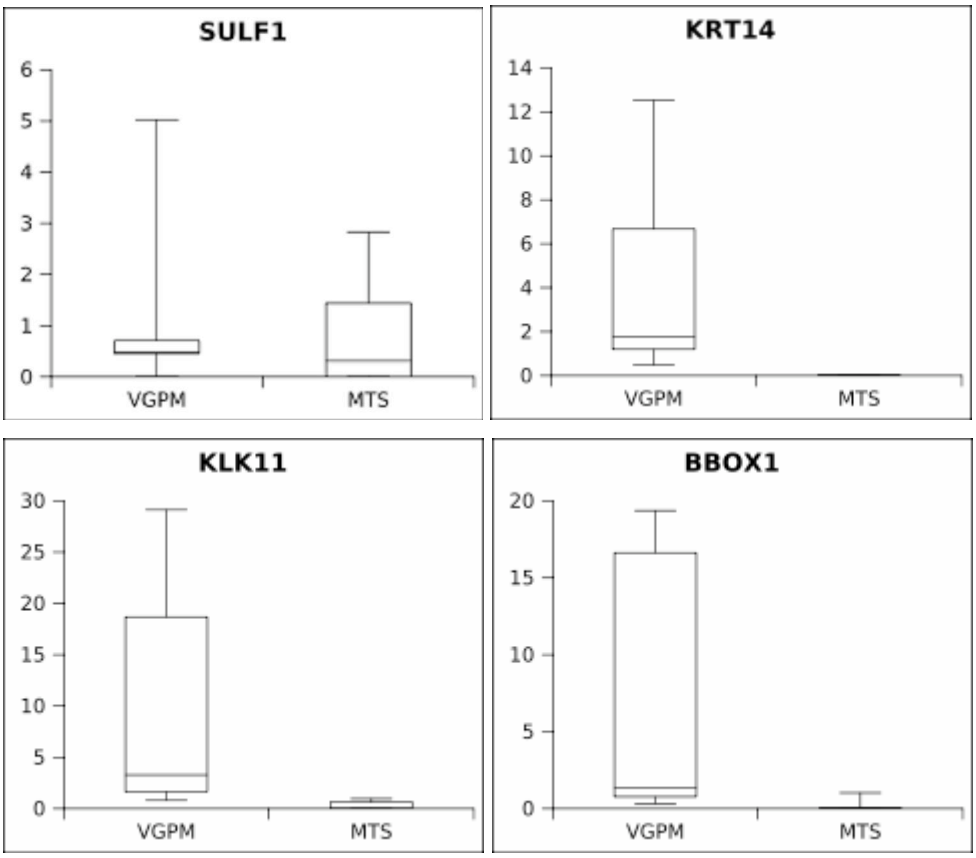

Figure 4: Box-plots comparing the qPCR results obtained for 5 independent samples from two series (MTS vs VGPM) for four genes (SULF1, KRT14, KLK11 and BBOX1).

Dysplastic nevi and progression trend

Finally, the questionable behaviour of dysplatic lesions was investigated by looking for differentially expressed transcripts in both two and multi-class comparison approaches.

The direct comparison between CMN and DN revealed only 24 differentially expressed genes implicated in cellular adhesion and neurogenesis.

When the same DN class was compared to the RGPM one, 100 deregulated transcripts were found, including some important oncogenes.

On the other hand, multiclass comparison by means of ANOVA enlighted that some of the genes upregulated in DN vs CMN behave more similarly to VGPM than to RGPM. This is the case of a set of Mismatch Repair System (MMR) transcripts.

A the same time, the class of dysplastic nevi seems highly heterogeneous, as demonstrated by the clustering results of Fig1, in which DN samples are found both within the left patient group, together with all but one RGPM, and within the group on the right,

with all the MTS and 10 VGPM. Such separation of DN does not correlate with their degree of atypia, but rather to the ratio between mitosis and apoptosis (higher in the DN subset co-clustered with VGPM), as revealed by morphological inspection (table VI). The group of genes that mainly distinguish the two clusters (bottom of Fig1) is enriched by cellular detoxification, RNA processing and antigen presentation biological processes.
Altogether, these results indicate that dysplatic nevi show heterogeneous features that do not characterize them as an intermediate step within melanoma progression.

Table VI: list of DN showing mitosis and apoptosis data. The upper class is composed by DN similar to RGPM, the bottom class represents DN similar to VGPM. The last column indicates the ration between mitosis and apoptosis for each samples and the average value for both class.

Code	MITOSIS	APOPTOSIS	M/A ratio
23	0.00	0.10	0.00
8b	0.00	0.20	0.00
24	0.10	0.30	0.33
19	0.00	0.20	0.00
21	0.00	0.10	0.00
18	0.00	0.10	0.00
26	0.20	0.20	1.00
class avergage			**0.19**
25a	0.10	0.10	1.00
25b	0.10	0.10	1.00
22	0.10	0.10	1.00
20	0.00	0.10	0.00
class avergage			**0.75**

Discussion

To date molecular profiling studies of melanoma have resulted into different and sometimes inconsistent data, due to lesion heterogeneity and different sample preparation techniques. In this work, for the first time, a whole genome expression profiling of melanocityc lesions, belonging to different melanoma progression stages, were analyzed and compared with a cohort of dysplastic nevi.

A major aim of this study was the identification of characteristic genes for each step of melanoma progression. Thus, a large-scale gene expression analysis was performed on common melanocytic and dysplastic nevi, radial and vertical growth phase melanomas and metastases.

At first, as model of melanomagenesis we analyzed the transition between CMN to RGPM. In this step only few genes were differentially expressed. Among these, HEY1, GDF15,

and ICAM5 were found consistently upregulated in RGPM, while PKIA and SORBS1 were overexpressed in CN. HEY1 (Hairy/enhancer-of-split related with YRPW motif 1) is a downstream effector of Notch signaling pathway. It is known that Notch1 signaling promotes primary melanoma progression by activating mitogen-activated protein kinase/phosphatidylinositol 3-kinase-Akt pathways and up-regulating N-cadherin expression (Liu ZJ, et al. Cancer Res. 2006). Accordingly, in a previous study (Talantov D,et al. Clin. Cancer. Res 2005) the gene GDF15 (Growth differentiation factor 15) was found associated with malignant melanoma but not with benign melanocytic lesions. SORBS1 is a gene potentially implicated in DNA repair via the activation of the proapoptotic pathway when the DNA damage is too severe to be repaired. The over expression of this gene in CN can explain how benign lesions do not evolve in tumoral lesions. The protein encoded by PKIA is a member of the cAMP-dependent protein kinase (PKA) inhibitor family. It has been shown (Becker D et al. Oncogene 1990) that in normal human melanocytes there is no expression of the catalytic or regulatory subunits of human PKA, whereas primary and metastatic melanomas express PKA. This can be explained by the upregulation of the PKA inhibitor (PKIA) in CN.

In this step we also confirmed the overexpression of AMICA1 and RNASE1, and the downmodulation of NUDT10 in RGPM compared to CMN. These three genes are poorly known: AMICA1 is an adhesion molecule that interacts with CXADR antigen 1, RNASE1 encodes a monomeric protein that acts to degrade ds-RNA over ss-RNA and NUDT10 belongs to a subgroup of phosphohydrolases that preferentially attack diphosphoinositol polyphosphates (Hidaka et al., 2002). At present, there are no information about the role of these genes in tumorigenesis.

From the comparison between RGPM and VGPM, several differentially expressed genes were found. Of note, the set of genes highly expressed in RGPM includes WNT3 and EGFR, while PTEN is downregulated. In a very recent paper, Castiglia (Castiglia D et al. Genes Chromosomes Cancer. 2008) affirms that the deregulation of the Wnt/beta-catenin signaling pathway cooperates to promote melanoma development and/or progression. In our cohort of samples WNT3 is more expressed in RGPM than in VGPM; this can indicate that WNT is more important for the early stages of melanoma progression than for advanced stages. Other important oncogenes such as NRAS, AKT and CTNNB1 were found overexpressed in VGPM. In a novel mouse model for melanoma (Delmas V et al. Genes Dev. 2007) it has been shown that stabilized beta-catenin bypasses the

requirement for p16(Ink4a) mutations and, together with an activated N-Ras oncogene, leads to melanoma with high penetrance and short latency. This data is in agreement with our results. So, there is a synergy between the Wnt and mitogen-activated protein (MAP) kinase pathways which may represent an important mechanism underpinning the genesis of melanoma. Some novel genes never described in literature were found differentially expressed between RGPM and VGPM. These include TNS3, PAIP1 and DAPL1, whose expression had also been validated by RT-qPCR. We also confirmed by RT-qPCR the overexpression of HSP90AB1 (Heat shock protein 90kDa alpha, class B member 1) in VGPM. This gene has been associated with cell motility and cancer metastasis; moreover, it is a marker of melanoma progression and also a drug target (McCarthy MM et al. Annals of Oncology 2008). The protein encoded by TNS3 (Tensin 3) is a positive regulator of cell motility (Cui Y et al. Mol. Can. Res. 2004), which makes sense with the overexpression of this gene in VGPM. No relevant functional information was found for the two other transcripts, PAIP1 (Poly(A) binding protein interacting protein 1) and DAPL1 (Death associated protein-like 1). PAIP1 is overexpressed in VGPM and acts as a coactivator in the regulation of translation initiation of poly(a)-containing mRNAs. We can hypothesize that transcription is more active in VGPM than in RGPM. In literature there is only one paper about DAPL1 in Prokaryotes (Hudson AO et al. J Bacteriol 2008). In analogy to DAP, we can suppose that also DAPL1 acts as a positive mediator of programmed cell death that is induced by interferon-gamma. In our samples this gene in downregulated in VGPM.

Concerning metastases, many significant and robust differences were found, regardless of the site and of the distance from primary tumor. During this transition some genes belonging to EGFR signaling pathway were downregulated (EGFR, AREG, EREG). This suggests that in advanced melanoma stages anti- EGFR therapy could be less effective, since it is known that response to Cetuximab correlates with the expression levels of EREG and AREG (Khambata-Ford S et al. J Clin Oncol. 2007).
For this step we confirmed the downmodulation in MTS of KLK11, BBOX1 and KRT14, as well as the weak overexpression of SULF1. KLK11 (Kallikrein-related peptidase 11) has been related to prostate, ovary and breast tumorigenesis but never associated with melanoma. Our results indicate that KLK11 is strongly downregulated in MTS and KLK5-7-8-10, too. The same trend is followed by BBOX1 (Butyrobetaine (gamma), 2-oxoglutarate dioxygenase) an enzyme responsible for the biosynthesis of l-carnitine, a key molecule of

fatty acid metabolism. Although the protein structure and activity have been extensively described, little information is available concerning BBOX1 expression.
The downregulation of 13 genes belonging to Keratins family observed in MTS is in accordance to Riker AI et al (BMC Medical Genomics, 2008) and Jaeger J et al. (Clin Cancer Res., 2007) and, together with the consistent downmodulation of keratins, collagens and laminins (table V) clearly demonstrate the acquisition of a higher cellular motility.

Among the highest deregulated genes between DN and CMN, a potent neurite outgrowth inhibitor (RTN4), was found less expressed in DN. On the other hand, HEY1, that encodes a nuclear protein, acting as a transcriptional repressor, had a higher expression in DN. This gene is induced by the Notch and c-Jun signal transduction pathways, but in contrast to HES1, another Notch1 canonical target, has never been associated to melanocytic lesions. We also found the upregulation of CD44, a cell-surface glycoprotein involved in cell-cell interactions, cell adhesion, migration and possibly related to tumor metastasis.
In the comparison between DN and RGPM, some important oncogenes (such as MYC and BCL6) were found less expressed in DN.
The DN class is highly heterogeneous. In particular, a group of genes enriched in cellular detoxification (MTHFR, CYP1B1, CYB5R1), RNA processing (DGCR8, DDEF1IT1, SMG1) and antigen presentation (B2M) processes behave differently within the DN class. They are upregulated in DN more similar to RGPM and downregulated in DN that co-cluster with VGPM and MTS. About RNA processing, the gene DGCR8 plays an important role in miRNA synthesis because the Drosha-DGCR8 complex initiates microRNA maturation by precise cleavage of the stem loops that are embedded in primary transcripts (pri-miRNAs). A downregulation of this gene can cause an inefficient miRNA synthesis. At the same time, the gene SMG1 encodes a protein involved in nonsense-mediated mRNA decay (NMD) as part of the mRNA surveillance complex. Also this transcript was found downregulated in DN similar to VGPM. As far as antigen presentation, a downregulation of Beta-2-Microglobulin (B2M) due to micro-deletion in the B2M gene, had also been found in the SK-MEL-33 melanoma cell line, as well as *in vivo*, demonstrating that this molecular lesion is likely to reflect a somatic mutation acquired during tumor progression (PMID: 8432869). Another difference detected between the two DN subsets is the average ratio between mitotic figures and apoptotic bodies, which is higher in DN similar to VGPM. This means that such subset has a proliferation advantage compared to RGPM-similar dysplastic nevi.

Taken together, all these results show that DN are similar to CMN to some extent, but also close to RGPM and VGPM as far as other molecular characteristics. For this reason, they should not be considered as an intermediate step within melanoma progression, but rather a separate although heterogeneous class with an independent risk to develop melanoma. A general upregulation of genes involved in MMR system was found in the transition from RGPM to VGPM and in DN. This implies that in both VGPM and DN the accumulation of DNA damage is sufficient to activate DNA mismatch repair mechanisms. Also in the transition from VGPM to MTS, mechanisms of DNA repair are activated, but in this case the differentially expressed genes are involved into DNA double-strand breaks repair system. This suggests that during melanoma progression, together with higher motility and proliferation, there is also the accumulation of a consistent DNA damage.

Acknowledgments

We thank Drs. Alessandro Zaccagna and Franco Picciotto of IRCC, Turin, for sample clinical collection. Drs. Luigi Bisi and Roberto Zanetti of CPO and Piedmont Cancer Registry, Turin, are acknowledged for the retrieval of follow-up information.