

# Replication of GWAS-identified neuroblastoma risk loci strengthens the role of *BARD1* and affirms the cumulative effect of genetic variations on disease susceptibility

Mario Capasso<sup>1,2,\*</sup>, Sharon J. Diskin<sup>3–5</sup>,  
Francesca Totaro<sup>1,2</sup>, Luca Longo<sup>6,7</sup>, Marilena  
De Mariano<sup>6,7</sup>, Roberta Russo<sup>1,2</sup>, Flora Cimmino<sup>1,2</sup>,  
Hakon Hakonarson<sup>4,8,9</sup>, Gian Paolo Tonini<sup>7</sup>, Marcella  
Devoto<sup>4,9–11</sup>, John M. Maris<sup>3–5</sup> and Achille Iolascon<sup>1,2</sup>

<sup>1</sup>Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli “Federico II”, Naples, Italy, <sup>2</sup>CEINGE - Biotecnologie Avanzate, Naples, Italy, <sup>3</sup>Division of Oncology and Center for Childhood Cancer Research, The Children’s Hospital of Philadelphia, Philadelphia, PA, <sup>4</sup>Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, <sup>5</sup>Abramson Family Cancer Research Institute, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, <sup>6</sup>Terapia Immunologica, IRCCS AOU San Martino-IST, National Cancer Research Institute, Genoa, Italy, <sup>7</sup>Italian Neuroblastoma Foundation, Genoa, Italy, <sup>8</sup>The Center for Applied Genomics, The Children’s Hospital of Philadelphia, Philadelphia, PA, <sup>9</sup>Division of Genetics, The Children’s Hospital of Philadelphia, Philadelphia, PA, <sup>10</sup>Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA and <sup>11</sup>University of Rome “La Sapienza”, Department of Molecular Medicine, Rome, Italy

\*To whom correspondence should be addressed. Tel: +390813737889;  
Fax: +390813737804;  
Email: [capasso@ceinge.unina.it](mailto:capasso@ceinge.unina.it)

Several neuroblastoma (NB) susceptibility loci have been identified within *LINC00340*, *BARD1*, *LMO1*, *DUSP12*, *HSD17B12*, *DDX4*, *IL31RA*, *HACE1* and *LIN28B* by genome-wide association (GWA) studies including European American individuals. To validate and comprehensively evaluate the impact of the identified NB variants on disease risk and phenotype, we analyzed 16 single nucleotide polymorphisms (SNPs) in an Italian population (370 cases and 809 controls). We assessed their regulatory activity on gene expression in lymphoblastoid (LCLs) and NB cell lines. We evaluated the cumulative effect of the independent loci on NB risk and high-risk phenotype development in Italian and European American (1627 cases and 2575 controls) populations. All NB susceptibility genes replicated in the Italian dataset except for *DDX4* and *IL31RA*, and the most significant SNP was rs6435862 in *BARD1* ( $P = 8.4 \times 10^{-15}$ ). *BARD1* showed an additional and independent SNP association (rs7585356). This variant influenced *BARD1* mRNA expression in LCLs and NB cell lines. No evidence of epistasis among the NB-associated variants was detected, whereas a cumulative effect of risk variants on NB risk (European Americans:  $P_{\text{trend}} = 6.9 \times 10^{-30}$ , Italians:  $P_{\text{trend}} = 8.55 \times 10^{-13}$ ) and development of high-risk phenotype (European Americans:  $P_{\text{trend}} = 6.9 \times 10^{-13}$ , Italians:  $P_{\text{trend}} = 2.2 \times 10^{-1}$ ) was observed in a dose-dependent manner. These results provide further evidence that the risk loci identified in GWA studies contribute to NB susceptibility in distinct populations and strengthen the role of *BARD1* as major genetic contributor to NB risk. This study shows that even in the absence of interaction the combination of several low-penetrance alleles has potential to distinguish subgroups of patients at different risks of developing NB.

## Introduction

Neuroblastoma (NB) is a neuroendocrine tumor, arising from any neural crest element of the sympathetic nervous system. Despite its

**Abbreviations:** CI, confidence interval; GWA, genome-wide association; LCLs, lymphoblastoid cell lines; LD, linkage disequilibrium; NB, neuroblastoma; OR, odds ratio; SNPs, single nucleotide polymorphisms.

relative low incidence accounting for <1000 cases in the USA and 150 cases in Italy per year, NB accounts for 15% of childhood cancer mortality (1). Currently, clinical trials stratify patients into four prognostic subgroups with expected very low risk, low risk, intermediate risk, and high risk of death from disease, and 16 pretreatment designations (2). Stratifying patients according to risk subgroups represent an important strategy to choose an appropriate and effective therapy. However, children with severe clinical course and widespread metastases categorized as ‘high risk’ have survival rates <35% despite aggressive and intensive therapies (1).

Recent genome-wide association studies (GWAS) demonstrated that common single nucleotide polymorphism (SNP) alleles were associated with NB. Some of these SNP associations were particularly enriched in the high-risk group for long intergenic non-protein coding RNA 340 (*LINC00340*) (3), BRCA1-associated RING domain 1 (*BARD1*) (4), and LIM domain only 1 (*LMO1*) (5), and in the low-risk group for dual specificity phosphatase 12 (*DUSP12*), hydroxysteroid (17-beta) dehydrogenase 12 (*HSD17B12*), DEAD (Asp-Glu-Ala-Asp) box polypeptide 4 (*DDX4*) and interleukin 31 receptor A (*IL31RA*) (6). Additional SNPs within HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1 (*HACE1*) and lin-28 homolog B (*C. elegans*) (*LIN28B*) have been found to be associated with NB but not with disease phenotype (7). These findings suggested that in NB common variants affect both tumor initiation and malignant progression.

All of the NB GWAS have been performed in European American population samples. Some associated SNPs at *LINC00340* (1), *BARD1* (4), and *LMO1* (5) have been replicated in a North European sample from the UK and in an African American sample (8). It is well known that European Americans form a structured population due to historical immigration from diverse source populations, and this can create false-positive genetic associations (9). Furthermore, among Europeans there is a consistent and reproducible distinction in SNP frequencies between “northern” and “southern” population groups (10). Based on these observations, it is evident that confirmatory studies are needed through replication in diverse populations to validate the GWAS findings.

In this study, we assessed the association of NB with eight SNPs (rs6939340 and rs4712653 located in *LINC00340*, and rs6435862, rs3768716, rs7585356, rs2070094, rs2229571 and rs1048108 located in *BARD1*) in a southern European population sample composed of 370 cases and 809 controls of Italian origin. In addition, we extended to this larger group the study of rs110419 and rs4758051 (*LMO1*), rs1027702 (*DUSP12*), rs11037575 (*HSD17B12*), rs2619046 (*DDX4*) and rs10055201 (*IL31RA*) already genotyped in a smaller sample (5,6). The results of the Italian case-control study were compared with those of the European American GWAS composed of 1627 cases and 2575 controls (6).

In our previous studies, we demonstrated that rs110419 in *LMO1* (5), rs6435862 in *BARD1* (11), and rs17065417 in *LIN28B* (7) are functional regulatory variants and are involved in NB tumorigenicity. In this study, we tested for SNP–gene expression associations in lymphoblastoid (LCL) and NB cell lines. Finally, we evaluated the cumulative effect of each GWAS identified and replicated genetic variants on NB risk and development of high-risk phenotype, and their potential interactions (epistasis), in both European American and Italian populations.

## Materials and methods

### Study subjects

This study consisted of 370 NB patients and 809 cancer-free controls of Italian origin. Case subjects were defined as children with a diagnosis of NB or ganglioneuroblastoma and collected through the Italian Neuroblastoma Group. The eligibility criterion for genotyping was the availability of DNA. All

control subjects were recruited from Italian blood donor centers. Eligibility criteria for control subjects were Italian origin, availability of DNA, no serious underlying medical disorder, including cancer.

In addition, this study included a GWAS dataset of 1627 NB patients registered through the North American-based Children's Oncology Group (COG) and 2575 cancer-free children of self-reported Caucasian ancestry who were recruited and genotyped by the Center for Applied Genomics at the Children's Hospital of Philadelphia (CHOP). European American cases and controls have been described in detail in a previous publication (6). Case subjects were defined as children with a diagnosis of NB or ganglioneuroblastoma and registered through the Children's Oncology Group. The blood samples from the patients with NB were identified through the NB biorepository of the Children's Oncology Group for specimen collection at the time of diagnosis. The eligibility criterion for genome-wide genotyping was the availability of 1.5 µg of DNA of high quality from a tumor-free source such as peripheral blood or bone marrow mononuclear cells that were uninvolved with a tumor. Control subjects were recruited from the Philadelphia region through the Children's Hospital of Philadelphia Health Care Network, including four primary care clinics and several group practices and outpatient practices that included well-child visits. Eligibility criteria for control subjects were European ancestry as determined by self-report or parental report, availability of 1.5 µg of high-quality DNA from peripheral-blood mononuclear cells, and no serious underlying medical disorder, including cancer.

Main clinical and biological characteristics of the patients, including age, stage of the disease (International Neuroblastoma Staging System), and v-myc myelocytomatosis viral related oncogene, neuroblastoma derived amplification status, are shown in Table I. Samples were assigned into two risk groups (not high risk and high risk) based on the Children's Oncology Group risk assignment (1). This study was approved by the Ethics Committee of the Medical University of Naples. The participants of this study gave written informed consent.

#### SNP selection and genotyping

We selected for genotyping 14 genetic variants that have previously been associated with NB in European American children. We included the two most significant SNPs identified in *LINC00340* (intronic rs6939340 and rs4712653) (1), in *BARD1* (intronic rs6435862 and rs3768716) (4) and in *LMO1* (intronic rs110419 and intergenic rs4758051) (5) as previously reported. In addition, we genotyped the putative functional variation rs7585356 located 3' downstream of *BARD1*, and its coding variants rs2070094, rs2229571 and rs1048108 (4). We also analyzed four SNPs (intergenic rs1027702 near *DUSP12*, intronic rs11037575 in *HSD17B12*, rs2619046 in *DDX4* and rs10055201 in *IL31RA*) recently found to be associated with low-risk NB (6). The Italian DNA samples were genotyped using SNP Genotyping Assay on 7900HT Real-time PCR system (Applied Biosystems). To monitor quality control, three DNA samples per genotype were genotyped by Sanger sequencing (3730 DNA analyzer, Applied Biosystems) and included in each 384-well reaction plate; genotype concordance was 100%. To confirm genotypes, we sequenced 20 samples chosen randomly from cases and controls; concordance between genotypes was 100%. Primer sequences are available on request. To test NB risk variants cumulative effect, we also used genotype data of rs4336470 (*HACE1*) and rs17065417 (*LIN28B*) already demonstrated to be associated with NB in this Italian case-control sample (7). The European American DNA samples were genotyped using the Illumina Infinium II BeadChip, HumanHap550 v1 and v3 array (Illumina) according to methods detailed elsewhere (3,4). The quality control analyses for this GWAS dataset are described in detail elsewhere (6).

#### SNP-gene expression correlation analysis

The influence of SNPs on gene expression was evaluated using two web tools, SNPExp v1.2 (<http://app3.titan.uio.no/biotools/tool.php?app=snpexp>) (12) and Genevar v3.1.0 (<http://www.sanger.ac.uk/resources/software/genevar/>) (13). SNPExp calculates correlation between HapMap genotypes and gene expression levels in LCLs using linear regression under an additive model. For this analysis, 198 unrelated HapMap3 subjects were chosen. In Genevar, a linear regression (under an additive model) was used to calculate the SNP-gene expression correlations performing a matched co-twin analysis. The expression and genotype data were obtained from LCLs of 170 individuals separate in two sets, Twin 1 (81 subjects) and Twin 2 (89 subjects). This analysis permits immediate replication and validation of the identified SNP-gene expression associations.

#### Genotyping of NB cell lines

For routine maintenance, cells were grown in RPMI 1640 (Sigma-Aldrich, R8758) complete media or Dulbecco's modified Eagle's medium (Sigma-Aldrich, D6546) complete media containing 10% fetal bovine serum (Sigma-Aldrich, F7524), 1X Penicillin-Streptomycin (Sigma-Aldrich, P0781) and 2mM L-Glutamine (Sigma-Aldrich, G7513). NB cell lines were genotyped for the *HSD17B12*-SNP rs11037575 and for the *BARD1*-SNP rs7585356. Genomic DNA was isolated using the Wizard Genomic DNA Purification Kit (Promega Corporation, Madison, WI). DNA (50ng) from NB cell lines was used as a template, and PCR

**Table I.** Clinical characteristic of study sample

Variable	Cases		P <sup>a</sup>
	Italians n = 370	European Americans n = 1627	
Age			
>=18 months	170 (46.8)	744 (52.9)	0.04
<18 months	193 (53.2)	662 (47.1)	
Unknown	7	221	
Sex			
Male	203 (56.1)	865 (53.2)	0.31
Female	159 (43.9)	762 (46.8)	
Unknown	8	—	
INSS Stage			
4	145 (40.2)	605 (43.2)	0.30
1,2,3,4s	216 (59.8)	795 (56.8)	
Unknown	9	227	
MYCN			
Amplified	80 (24.0)	233 (17.3)	0.005
Not amplified	254 (76.0)	1111 (82.7)	
Unknown	36	283	
Risk			
High risk	212 (57.3)	595 (43.7)	0.000003
Not high risk	158 (42.7)	767 (56.3)	
Unknown	—	265	
	Controls		
	n = 809	n = 2572	
Variable	Number (%)	Number (%)	
Sex			
Male	458 (56.6)	1325 (51.5)	0.01
Female	351 (42.3)	1247 (48.4)	
Age (mean years; SD) <sup>b</sup>	10.1 (6.4)	9.4 (5.6)	0.002

<sup>a</sup>Statistical comparison of clinical variables between the two case and control groups.

<sup>b</sup>Age data were not available for 359 European Americans

amplicons were generated using Promega PCR Master Mix. The DNA primer sequences were 'rs11037575' forward: 5'-GGTGGGTGTCATTCTTTT-3'; reverse: 5'-TACTTGGCTTTGGCCCATTA-3'; and 'rs7585356' forward: 5'-GTTCATGAGGAACCAACTGG-3'; reverse: 5'-TGGAGGCAGAAGTTG GTGAT-3'. Each amplicon was isolated on 1% agarose gel and purified using QIAamp purification kit (Qiagen). Finally, each amplicon was sequenced using 3730 DNA analyzer (Applied Biosystems). The NB cell lines and genotypes are reported in the Supplementary Table 1, available at *Carcinogenesis* Online.

#### Gene expression of NB cell lines

RNA was isolated from NB cell lines using the TRIzol reagent (Invitrogen Life Technologies). Two hundred nanograms of total RNA was reverse transcribed into cDNA using iScript cDNA Synthesis Kit (Bio-Rad). Quantitative RT-PCR (qRT-PCR) using Power SYBR Green PCR Master Mix (Applied Biosystems) was performed to evaluate the gene expression of *HSD17B12* and *BARD1*. Samples were amplified on an Applied Biosystems 7900HT Sequence Detection System using standard cycling conditions and data collected and analyzed by 2<sup>-ΔΔCt</sup> method as described in our previous article (14). β-Actin gene was used as housekeeping gene. Primers overlapping the exon-exon junction were used. Primer sequences were as follows: '*HSD17B12*' forward: 5'-TTGTAGATTTCTTCTCAGTG-3'; reverse: 5'-GCAGGACACTCTGCACAAA-3'; '*BARD1*' forward: 5'-GAC AACTGGACAGCATGATCAA-3'; reverse: 5'-TTGTTTCCTGCATCATTAAACAAAC-3'. Primers used for *HSD17B12* and *BARD1* were specifically designed to detect the cDNA full length isoforms.

#### Genotype imputation on chromosome 2q35 data

As SNPs rs7585356, rs2070094, rs2229571 and rs1048108 of *BARD1* were not included in Illumina HumanHap550 arrays, genotype imputation was performed in the European American GWAS dataset using PLINK software v1.06 (15). The HapMap genotypes (release 23) of CEU subjects, downloaded from the HapMap database (<http://www.hapmap.org>), were used as reference. A total of 1071 SNPs of the 2q35 band included in the arrays were used to perform the imputation analysis.

#### Statistical analysis

Hardy-Weinberg equilibrium was evaluated using the goodness-of-fit chi-square test in control subjects. Two-sided chi-square tests were used to

evaluate differences in the distributions of allele frequencies between all patients and controls, and only high-risk patients and controls. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated to assess the relative disease risk conferred by a specific allele. We defined replication as a  $P \leq 0.05$  with a consistent direction of association. We did not correct for multiple tests because the analyzed genes were previously reported associated to NB in multiple independent datasets (3–6,8). Student's  $t$ -test was used to compare the differences in the mRNA expression levels between NB cell lines stratified according to genotype. Linkage disequilibrium (LD) and haplotype analysis were performed using the web-site tool SNAP v2.2 (<http://www.broadinstitute.org/mpg/snap/index.php>) (16) and Haploview v4.2 software (17). We assessed the cumulative effects of the independent significant risk loci on NB initiation and phenotype, by counting the number of risk alleles in each subject and modeling the summary variable categorically in logistic regression analysis. The alleles with higher frequency in cases than controls for each SNP were defined as 'risk alleles', and genotypes were coded as 0, 1 or 2 according to the number of risk alleles. Cumulative risk scores were calculated for all SNPs and patients were grouped into categories based on the number of risk alleles 0–7, 8, 9, 10, 11–15 for case-control comparison, and 0–3, 4–5, 6–8 for case-only (high-risk versus not high-risk phenotype comparison). The risk categories were created keeping similar percentage of individuals in each group. For the case-only comparison of high-risk against not high-risk phenotype, we created only three risk categories to have more statistical power as the number of SNPs and patients was lower than in the case-control group. ORs were calculated comparing the groups defined by varying number of risk alleles to the group with the lowest number of risk alleles (0–7 for the case-control analysis and 0–3 for the case-only analysis). A  $P$  value of  $\leq 0.05$  was considered as statistically significant. This analysis was performed using IBM SPSS 19.0 software. Gene-gene interaction was tested using logistic regression additive model by adding an interaction term between the genotypes from all pair of SNPs investigated. This analysis was performed by PLINK software v1.06 (15).

## Results

Table I shows the clinical parameters of 370 and 1627 NB patients with Italian and European American origin, respectively. The allele frequencies of all selected genetic variants among cases and controls and their association with NB and high-risk phenotype are shown in Table II and Supplementary Table 2, available at *Carcinogenesis* Online. Because SNPs rs7585356, rs2070094, rs2229571 and rs1048108 were not included in the European American study, we performed genotype imputation at 2q35 (Supplementary Table 3, available at *Carcinogenesis* Online). The most significant imputed SNP was rs17487827, which is in strong LD with genotyped markers rs3768716 and rs6435862 ( $r^2 = 0.95$  and  $0.60$ , respectively in 1000 Genomes Project data). The observed genotype frequencies among the control subjects were in agreement with Hardy–Weinberg equilibrium, except for rs2070094 ( $P < 0.001$ ) in Italian children. This SNP was excluded from further analyses. Table III summarizes the characteristics of the studied SNPs.

In the Italian case-control study, significant differences of allele distributions were observed for all SNPs except rs4758051 (*LMO1*), rs2619046 (*DDX4*) and rs10055201 (*IL31RA*). SNP rs6435862 within *BARD1* was the most significant (Table II,  $P = 8.4 \times 10^{-15}$ ). All SNPs showed a similar pattern of association to NB in Italian and European American children. All SNPs of *LINC00340* and *BARD1*, and one of *LMO1* were significantly associated with high-risk NB in both populations (Table II and Supplementary Table 2, available at *Carcinogenesis* Online).

We then tested for SNP–gene expression associations at all 14 SNPs and at 2 previously reported SNPs (7). The analysis of gene expression variation using genome-wide expression arrays from LCLs of 198 unrelated HapMap individuals demonstrated that SNPs rs7585356 (*BARD1*),

**Table II.** Association of 16 previously identified SNPs with neuroblastoma in Italian children

SNP ID	Allele <sup>a</sup> Italians		SNP–gene expression associations											
	Case genotypes	Control genotypes	Case MAF	High-risk genotypes	Control genotypes	High-risk MAF	Control MAF	$P^b$	OR <sup>b</sup> (95% CI)	$P^c$	OR <sup>c</sup> (95% CI)	P HapMap <sup>d</sup>	P Twin1 <sup>e</sup>	P Twin2 <sup>e</sup>
<i>LINC00340</i> (rs6939340)	A/G	196/390/175	0.54	297/1148	196/390/175	0.56	0.49	0.01	1.25 (1.04–1.50)	0.01	1.37 (1.06–1.76)	N/A	N/A	N/A
<i>LINC00340</i> (rs4712653)	T/C	199/376/197	0.54	316/257	199/376/197	0.59	0.50	0.05	1.20 (1.00–1.44)	$5.3 \times 10^{-3}$	1.43 (1.11–1.83)	N/A	N/A	N/A
<i>BARD1</i> (rs6435862)	T/G	423/312/49	0.43	49/67/33	423/312/49	0.45	0.26	$8.4 \times 10^{-15}$	2.10 (1.73–2.53)	$1.2 \times 10^{-10}$	2.28 (1.76–2.94)	0.50	$4.8 \times 10^{-6}$	$3.0 \times 10^{-4}$
<i>BARD1</i> (rs768716)	A/G	473/289/45	0.35	60/74/20	473/289/45	0.37	0.23	$2.9 \times 10^{-9}$	1.78 (1.50–2.16)	$6.3 \times 10^{-7}$	1.91 (1.48–2.48)	0.16	$5.0 \times 10^{-4}$	$4.0 \times 10^{-4}$
<i>BARD1</i> (rs7585356)	G/A	370/305/94	0.23	94/51/7	370/305/94	0.21	0.32	$2.0 \times 10^{-5}$	0.63 (0.51–0.78)	$2.2 \times 10^{-4}$	0.57 (0.43–0.77)	$3.0 \times 10^{-4}$	$1.0 \times 10^{-4}$	$7.7 \times 10^{-8}$
<i>BARD1</i> (rs2070094)	G/A	419/247/109	0.20	103/33/9	419/247/109	0.18	0.30	$1.8 \times 10^{-6}$	0.59 (0.47–0.73)	$2.0 \times 10^{-5}$	0.50 (0.36–0.68)	N/A	N/A	N/A
<i>BARD1</i> (rs2229571)	G/C	238/363/208	0.35	74/53/24	238/363/208	0.33	0.48	$7.1 \times 10^{-9}$	0.58 (0.49–0.70)	$2.5 \times 10^{-6}$	0.54 (0.42–0.70)	N/A	$1.0 \times 10^{-3}$	$6.8 \times 10^{-6}$
<i>BARD1</i> (rs1048108)	C/T	296/355/99	0.27	69/39/10	296/355/99	0.25	0.37	$5.0 \times 10^{-5}$	0.64 (0.52–0.79)	$3.9 \times 10^{-4}$	0.57 (0.42–0.78)	N/A	N/A	N/A
<i>LMO1</i> (rs110419)	G/A	271/370/133	0.50	36/62/39	271/370/133	0.51	0.41	$5.0 \times 10^{-5}$	1.46 (1.21–1.76)	$2.0 \times 10^{-3}$	1.50 (1.16–1.94)	0.24	0.03	0.60
<i>LMO1</i> (rs758051)	A/G	246/405/141	0.44	47/72/23	246/405/141	0.42	0.43	0.94	1.00 (0.84–1.21)	0.57	0.93 (0.72–1.20)	0.34	0.20	0.22
<i>DUSP12</i> (rs1027702)	C/T	258/385/128	0.37	61/67/22	258/385/128	0.37	0.42	0.05	0.83 (0.69–1.00)	0.14	0.83 (0.64–1.06)	$4.0 \times 10^{-7}$	0.42	0.27
<i>HSD17B12</i> (rs11037575)	T/C	322/379/84	0.40	52/74/21	322/379/84	0.39	0.35	0.03	1.23 (1.02–1.47)	0.13	1.22 (0.94–1.57)	$1.1 \times 10^{-8}$	$7.2 \times 10^{-6}$	$2.0 \times 10^{-4}$
<i>DDX4</i> (rs2619046)	G/A	456/300/44	0.27	78/60/8	456/300/44	0.26	0.24	0.19	1.15 (0.93–1.41)	0.52	1.10 (0.83–1.46)	0.30	0.24	0.69
<i>IL31RA</i> (rs10055201)	A/G	475/264/36	0.24	90/54/7	475/264/36	0.23	0.22	0.26	1.13 (0.91–1.40)	0.75	1.05 (0.78–1.41)	0.68	0.26	0.41
<i>HACE1</i> (rs4336470)	C/T	348/329/103	0.30 <sup>f</sup>	63/68/18	348/329/103	0.35	0.34	<sup>h</sup> 0.06	<sup>h</sup> 0.71 (0.52–0.97)	0.84	1.03 (0.79–1.33)	0.02	0.23	0.60
<i>LINC28B</i> (rs17065417)	A/C	600/155/8	0.08 <sup>f</sup>	126/22/2	600/155/8	0.09	0.11	<sup>h</sup> 0.03	<sup>h</sup> 0.83 (0.68–1.00)	0.20	0.75 (0.49–1.16)	N/A	0.92	0.14

MAF, minor allele frequency.

<sup>a</sup>Major/minor Alleles; Italians: 370 cases and 809 controls.

<sup>b</sup> $P$  values and ORs from comparison of case versus control.

<sup>c</sup> $P$  values and ORs from comparison of high risk versus control.

<sup>d</sup> $P$  values from SNP–gene expression correlation using 198 unrelated HapMap subjects (SNPexp web tool).

<sup>e</sup> $P$  values from SNP–gene expression correlation using 170 subjects (Genevar web tool).

<sup>f</sup>Data as reported in Diskin *et al.* 2012.

N/A: Not available.

**Table III.** Features of analyzed SNPs

n	SNP ID	Chromosome band	Function	Allele <sup>a</sup>	Italian dataset SNP type	EA dataset SNP type
1	<i>LINC00340</i> (rs6939340)	6p22	—	A/G	Genotyped	Genotyped
2	<i>LINC00340</i> (rs4712653)	6p22	—	T/C	Genotyped	Genotyped
3	<i>BARD1</i> (rs6435862)	2q35	Intronic	T/G	Genotyped	Genotyped
4	<i>BARD1</i> (rs3768716)	2q35	Intronic	A/G	Genotyped	Genotyped
5	<i>BARD1</i> (rs7585356)	2q35	3'downstream	G/A	Genotyped	Imputed
6	<i>BARD1</i> <sup>b</sup> (rs2070094)	2q35	Exon 6 (Val507Met)	G/A	Genotyped	—
7	<i>BARD1</i> (rs2229571)	2q35	Exon 4 (Arg378Ser)	G/C	Genotyped	Imputed
8	<i>BARD1</i> (rs1048108)	2q35	Exon 1 (Pro24Ser)	C/T	Genotyped	Imputed
9	<i>LMO1</i> (rs110419)	11p15	Intronic	G/A	Genotyped	Genotyped
10	<i>LMO1</i> (rs4758051)	11p15	Intergenic	A/G	Genotyped	Genotyped
11	<i>DUSP12</i> (rs1027702)	1q23.3	Intergenic	C/T	Genotyped	Genotyped
12	<i>HSD17B12</i> (rs11037575)	11p11.2	Intronic	T/C	Genotyped	Genotyped
13	<i>DDX4</i> (rs2619046)	5p15.2-p13.1	Intronic	G/A	Genotyped	Genotyped
14	<i>IL31RA</i> (rs10055201)	5p15.2-p13.1	Intronic	A/G	Genotyped	Genotyped
15	<i>HACE1</i> (rs4336470)	6q16.3	Intronic	C/T	Genotyped <sup>c</sup>	Genotyped
16	<i>LIN28B</i> (rs17065417)	6q21	Intronic	A/C	Genotyped <sup>c</sup>	Genotyped

<sup>a</sup>Major/minor alleles. <sup>b</sup>Excluded as not in agreement with Hardy–Weinberg equilibrium.

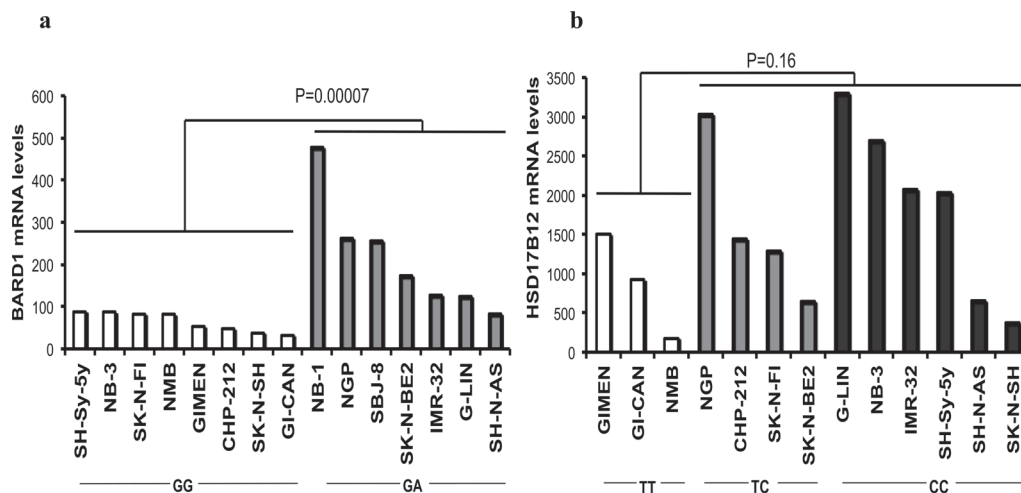
<sup>c</sup>Already genotyped in Diskin et al. 2012.

rs1027702 (*DUSP12*), rs4336470 (*HACE1*) and rs11037575 (*HSD17B12*) affected expression of the respective genes (Table II). These results were confirmed in 170 additional LCLs only for rs7585356 (*BARD1*) and rs11037575 (*HSD17B12*) when we performed a matched co-twin analysis, which permits immediate replication and validation of expression quantitative trait loci discoveries (Table II, Supplementary Figure 1, available at *Carcinogenesis* Online). These analyses demonstrated a high reproducibility of results for these two putative regulatory variants, whereas the other SNPs did not show the same results in different populations. Therefore, a qRT–PCR gene expression analysis was performed only for rs7585356 and rs11037575. The mRNA expression of full length *BARD1* isoform was significantly higher in NB cell lines heterozygous at rs7585356 (AG) (Figure 1a). A trend toward association between high mRNA levels of full length *HSD17B12* isoform and presence of the risk allele C was observed for rs11037575 without reaching the threshold for statistical significance (Figure 1b).

To select SNPs independently associated with NB among the 13 replicated ones, we performed a LD analysis using data from 1000 Genomes Project (<http://www.broadinstitute.org/mpg/snap/index.php>). SNPs with a  $r^2 > 0.10$  with the most significant SNP within each gene were removed (Supplementary Table 4, available at *Carcinogenesis* Online). Consequently, eight SNPs (rs6939340

(*LINC00340*), rs6435862 and rs7585356 (*BARD1*), rs110419 (*LMO1*), rs1027702 (*DUSP12*), rs11037575 (*HSD17B12*), rs4336470 (*HACE1*) and rs17065417 (*LIN28B*) were selected for further analyses. Of note, the LD and haplotype analyses demonstrated that SNPs rs6435862 and rs7585356 of *BARD1* were located in two independent genetic loci (Supplementary Figure 2, available at *Carcinogenesis* Online). To confirm the independence of their association with NB risk, we conducted a logistic regression analysis by adjusting for either of the two SNPs, and found that both SNPs still remained significant in European Americans (rs6435862:  $P = 1.8 \times 10^{-10}$ , OR = 1.25, 95% CI: 1.11–1.40; rs7585356:  $P = 0.04$ , OR = 0.88, 95% CI: 0.78–0.99) and Italians (rs6435862:  $P = 1.7 \times 10^{-4}$ , OR = 1.97, 95% CI: 1.60–2.43; rs7585356:  $P = 0.04$ , OR = 0.80, 95% CI: 0.64–0.99).

The association of these eight independent risk variants and NB stratified by clinical stage (stage 4 versus not stage 4), MYCN status (amplified versus not amplified), risk assessment (high risk versus not high risk), and age at diagnosis (age  $\geq 18$  months and  $< 18$  months) was further evaluated among Italian and European American cases. As shown in Table IV, in the European American sample the genetic variants of *LINC00340* (3), *BARD1* (4) and *LMO1* (5) confirmed the association with more clinically aggressive phenotypes, as already demonstrated in the previous articles. In the Italian patients, these SNPs showed the same trend for association, but only rs7585356

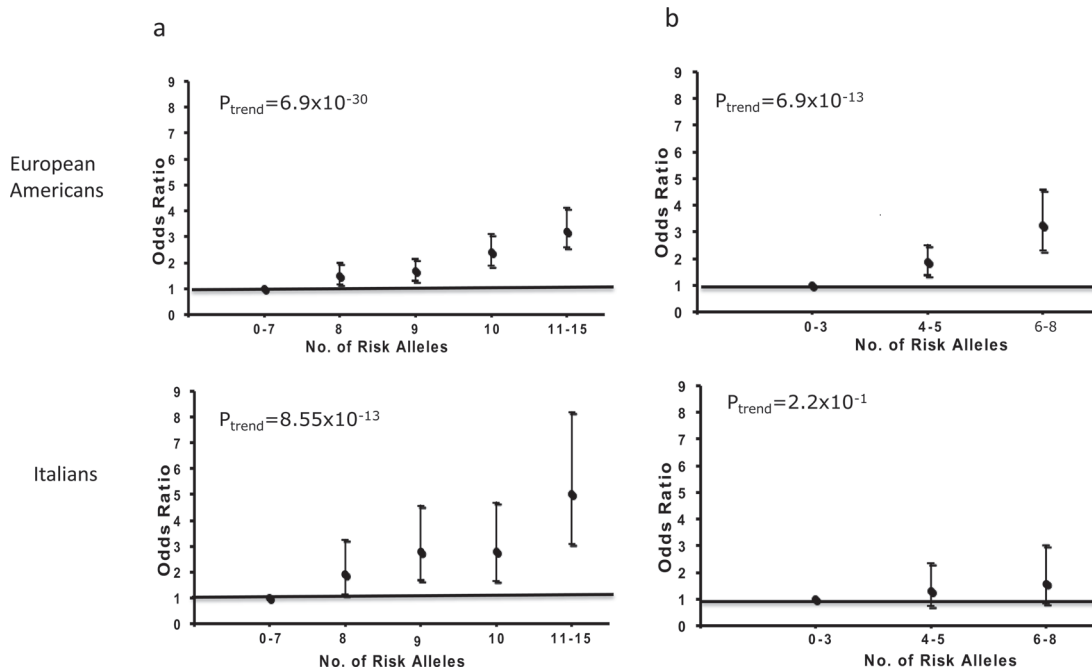


**Fig. 1.** qRT–PCR analysis to test the SNP–gene expression correlation in neuroblastoma (NB) cell lines. (a) mRNA expression of full length *BARD1* isoform in NB cell lines stratified according to rs7585356 genotype. No NB cell lines carried the genotype AA. (b) mRNA expression of full length *HSD17B12* isoform in NB cell lines stratified according to rs11037575 genotype.

**Table IV.** Association of the eight genetic variants with pathologic characteristics of neuroblastoma

SNP ID	Stage 4 N (%)	Not Stage 4 N (%)	P	OR (95% CI)	Stage 4 N (%)	Not Stage 4 N (%)	P	OR (95% CI)	MYCN Amp N (%)	Not MYCN Amp N (%)	P	OR (95% CI)	MYCN Amp N (%)	Not MYCN Amp N (%)	P	OR (95% CI)
<i>LINC00340</i> (rs6939340)	155 (0.56)	206 (0.54)	0.52	1.11 (0.81–1.51)	561 (0.60)	648 (0.54)	9.0x10 <sup>-4</sup>	1.29 (1.11–1.51)	89 (0.57)	245 (0.54)	0.44	1.15 (0.80–1.67)	226 (0.62)	924 (0.54)	2.0 x 10 <sup>-3</sup>	1.38 (1.13–1.70)
<i>BARD1</i> (rs6435862)	119 (0.43)	169 (0.43)	0.95	1.01 (0.74–1.38)	516 (0.40)	358 (0.30)	7.2x10 <sup>-8</sup>	1.54 (1.31–1.80)	75 (0.47)	187 (0.41)	0.17	1.29 (0.90–1.85)	170 (0.41)	644 (0.33)	1.7 x 10 <sup>-3</sup>	1.39 (1.13–1.70)
<i>BARD1</i> <sup>a</sup> (rs7585356)	52 (0.19)	98 (0.25)	0.04	0.67 (0.46–0.98)	729 (0.25)	845 (0.32)	2.6x10 <sup>-4</sup>	0.71 (0.58–0.85)	35 (0.23)	106 (0.23)	0.92	1.02 (0.66–1.58)	293 (0.22)	1205 (0.31)	8.3 x 10 <sup>-4</sup>	0.64 (0.49–0.83)
<i>LMO1</i> (rs110419)	131 (0.51)	187 (0.50)	0.98	1.01 (0.73–1.38)	576 (0.59)	654 (0.54)	2.8x10 <sup>-3</sup>	1.26 (1.46–1.08)	71 (0.49)	230 (0.52)	0.60	0.91 (0.62–1.32)	208 (0.55)	961 (0.56)	0.58	0.94 (0.77–1.16)
<i>DUSP12</i> (rs1027702)	104 (0.38)	154 (0.38)	0.99	0.99 (0.73–1.37)	528 (0.34)	413 (0.33)	0.56	1.05 (0.89–1.23)	59 (0.37)	179 (0.38)	0.81	0.96 (0.66–1.39)	749 (0.36)	168 (0.33)	0.24	1.13 (0.92–1.40)
<i>HSD17B12</i> (rs11037575)	106 (0.38)	159 (0.40)	0.56	0.91 (0.66–1.25)	753 (0.43)	524 (0.48)	0.03	0.84 (0.72–0.98)	56 (0.36)	197 (0.42)	0.25	0.80 (0.55–1.17)	1042 (0.44)	206 (0.47)	0.29	0.90 (0.73–1.10)
<i>HACE1</i> (rs4336470)	91 (0.32)	116 (0.29)	0.31	1.19 (0.85–1.65)	484 (0.29)	351 (0.31)	0.33	0.92 (0.78–1.09)	62 (0.42)	135 (0.28)	0.00	1.92 (1.31–2.82)	667 (0.30)	138 (0.30)	0.78	0.97 (0.78–1.21)
<i>LINC28B</i> (rs17065417)	22 (0.08)	35 (0.09)	0.78	0.92 (0.53–1.61)	133 (0.08)	92 (0.08)	0.46	0.90 (0.68–1.19)	19 (0.13)	36 (0.07)	0.04	1.83 (1.02–3.30)	169 (0.09)	43 (0.08)	0.24	1.23 (0.87–1.75)
SNP ID	High risk, n (%)	Not high risk, n (%)	P	OR (95% CI)	High risk, n (%)	Not high risk, n (%)	P	OR (95% CI)	Age >= 18 months, n (%)	Age < 18 months, n (%)	P	OR (95% CI)	Age >= 18 months, n (%)	Age < 18 months, n (%)	P	OR (95% CI)
<i>LINC00340</i> (rs6939340)	176 (0.55)	192 (0.53)	0.60	1.08 (0.80–1.47)	556 (0.60)	622 (0.54)	1.3x10 <sup>-3</sup>	1.29 (1.10–1.50)	160 (0.55)	203 (0.55)	0.91	1.02 (0.75–1.39)	677 (0.58)	540 (0.54)	0.01	1.21 (1.04–1.40)
<i>BARD1</i> (rs6435862)	142 (0.44)	151 (0.41)	0.45	1.12 (0.83–1.52)	439 (0.41)	410 (0.31)	2.1x10 <sup>-7</sup>	1.52 (1.30–1.78)	134 (0.45)	155 (0.41)	0.24	1.20 (0.88–1.64)	516 (0.39)	358 (0.31)	6.4 x 10 <sup>-5</sup>	1.37 (1.18–1.61)
<i>BARD1</i> <sup>a</sup> (rs7585356)	65 (0.21)	93 (0.24)	0.36	0.84 (0.59–1.21)	722 (0.24)	817 (0.32)	1.4x10 <sup>-4</sup>	0.69 (0.57–0.84)	59 (0.19)	93 (0.26)	0.02	0.64 (0.44–0.92)	873 (0.27)	704 (0.32)	5.2 x 10 <sup>-3</sup>	0.77 (0.64–0.92)
<i>LMO1</i> (rs110419)	150 (0.51)	176 (0.50)	0.92	1.02 (0.74–1.38)	575 (0.60)	628 (0.53)	5.9x10 <sup>-4</sup>	1.31 (1.12–1.52)	140 (0.52)	179 (0.49)	0.25	1.20 (0.88–1.65)	694 (0.59)	539 (0.53)	4.31 x 10 <sup>-3</sup>	1.24 (1.07–1.44)
<i>DUSP12</i> (rs1027702)	123 (0.38)	137 (0.37)	0.72	1.06 (0.78–1.44)	517 (0.34)	403 (0.34)	0.86	1.01 (0.86–1.19)	116 (0.38)	142 (0.34)	0.87	1.03 (0.75–1.40)	438 (0.34)	507 (0.33)	0.58	1.04 (0.89–1.22)
<i>HSD17B12</i> (rs11037575)	121 (0.38)	151 (0.41)	0.36	0.87 (0.64–1.18)	717 (0.44)	526 (0.47)	0.18	0.90 (0.77–1.05)	115 (0.38)	153 (0.41)	0.57	0.91 (0.67–1.25)	611 (0.46)	675 (0.46)	0.75	0.97 (0.84–1.13)
<i>HACE1</i> (rs4336470)	104 (0.35)	108 (0.27)	0.02	1.46 (1.05–2.02)	455 (0.30)	351 (0.30)	0.89	0.98 (0.84–1.17)	104 (0.32)	103 (0.28)	0.30	1.19 (0.86–1.65)	404 (0.30)	440 (0.31)	0.66	0.96 (0.82–1.13)
<i>LINC28B</i> (rs17065417)	26 (0.09)	32 (0.08)	0.74	1.10 (0.64–1.88)	127 (0.07)	89 (0.08)	0.44	0.89 (0.67–1.18)	27 (0.08)	30 (0.08)	0.89	1.04 (0.60–1.79)	105 (0.08)	120 (0.08)	0.90	1.02 (0.77–1.34)

Data from Italian population are colored in grey.  
<sup>a</sup>Imputed in the US population.



**Fig. 2.** Odds ratio for (a) neuroblastoma patients and (b) high-risk patients according to increasing number of risk alleles in Italian and European American cases for the identified risk loci (a) rs6939340, rs6435862, rs7585356, rs110419, rs1027702, rs11037575, rs4336470 and rs17065417; (b) rs6939340, rs6435862, rs7585356 and rs110419). The vertical bars represent 95% confidence intervals. The horizontal line denotes the null value (OR = 1).

resulted to be significantly associated with stage 4 and age  $\geq 18$  months ( $P = 0.04$ , OR = 0.67, 95% CI = 0.46–0.98 and  $P = 0.02$ , OR = 0.64, 95% CI = 0.44–0.92).

We evaluated potential pairwise interaction effects among the studied SNPs. No evidence of epistasis was detected through this analysis in both populations (Supplementary Table 5, available at *Carcinogenesis* Online). We further investigated the cumulative effects of the eight susceptibility loci on NB development in both populations using as predictor the total number of risk alleles carried by individual subjects. In European Americans, the distribution of risk alleles carried in both cases and controls followed a normal distribution, but with a shift toward a higher number of risk alleles in the cases. This result was confirmed in Italian population (Supplementary Figure 3a, available at *Carcinogenesis* Online). As shown in Figure 2a, European American individuals with multiple risk alleles had higher risk of developing NB compared with those with 0–7 risk alleles of the eight variants, in a dose-dependent manner with increasing number of risk alleles ( $P_{\text{trend}} = 6.9 \times 10^{-30}$ , OR = 1.23, 95% CI: 1.19–1.28). We were able to confirm this result in Italian ( $P_{\text{trend}} = 8.55 \times 10^{-13}$ , OR = 1.35, 95% CI: 1.24–1.46). We also tested for the cumulative effects of the four genetic variants (rs6939340, rs6435862, rs7585356 and rs110419) significantly found to be associated with clinically aggressive NB subgroups in European Americans (Table IV). The distribution of risk alleles carried in both not high-risk and high-risk patients is shown in Supplementary Figure 3b, available at *Carcinogenesis* Online. A shift toward a higher number of risk alleles in high-risk individuals was observed in European Americans but it was less evident in Italians. As shown in Figure 2b, the OR relative to the 0–3 risk alleles group significantly increased with increasing number of risk alleles in European Americans ( $P_{\text{trend}} = 6.9 \times 10^{-13}$ , OR = 1.38, 95% CI: 1.26–1.51). The same trend was confirmed in Italians without reaching the threshold for statistical significance ( $P_{\text{trend}} = 2.2 \times 10^{-1}$ , OR = 1.11, 95% CI: 0.94–1.30).

## Discussion

In this study, we set out to replicate a number of GWAS-identified NB susceptibility loci in Italian NB patients and healthy controls. All SNPs at the *BARD1* locus showed a strong association, whereas the association

with SNPs in *LINC00340*, *LMO1*, *DUSP12*, *HSD17B12*, *HACE1* and *LIN28B* was more moderate. SNPs at the *DDX4* and *IL31RA* locus showed a trend toward association with NB, but did not reach statistical significance. Interestingly, SNP rs7585356 located 3' downstream of *BARD1* was found to be an additional and independent risk factor for NB. These findings strengthen the role of *BARD1* as NB susceptibility gene. Indeed, we have already demonstrated that the genetic variant rs6435862 was strongly associated with high-risk NB (4) and correlated with high expression of the oncogenic *BARD1*β isoform (lacking exons 2 and 3), that led to an increased tumorigenicity of NB cell lines (11). Recently, NB SNP association at the *BARD1* locus has been replicated in African Americans, whereas limited association has been found at *LINC00340* and *LMO1* in the same population (8). In this study, we have demonstrated that rs7585356, independently from rs6435862, was associated with NB and also influenced full length *BARD1* isoform expression in NB cell lines. Based on these data, it is evident that more than one disease-contributing *BARD1* variant may exist. Of note, the less frequent genotype of SNP rs7585356 showed a protective role in NB development and a correlation with increased expression of the full length *BARD1* isoform. This suggests an oncosuppressor role of *BARD1* in the biology of neuroblastic malignant transformation, as also recently indicated for colon tumorigenesis (18). Given the complex structure of *BARD1* that displays diverse domains (RING, ANK and BRCT), this gene has been indicated to have multiple functions (19). These functions might be regulated by the expression of differentially spliced isoforms. So, it is reasonable to hypothesize that common and rare variants might affect the expression of distinct *BARD1* isoforms that in turn might have tumor suppression and oncogenic functions in NB. Further studies are needed to address this hypothesis.

Our genetic association analysis of clinical NB phenotypes confirmed that four SNPs at *LINC00340* (rs6939340), *BARD1* (rs6435862 and rs7585356) and *LMO1* (rs110419) conferred an increased risk for high-risk phenotype in European American patients. In Italian patients, only *BARD1* SNP rs7585356 showed a significant association with advanced INSS stage and age older than 18 months. We hypothesize that this lack of association to the other SNPs is due to insufficient statistical power because of the limited sample size. Indeed, the trend toward association with clinically aggressive phenotypes was the same for all four variants in European American and Italian patients.

This article shows that eight independent NB-associated common genetic variants (rs6939340, rs6435862, rs7585356, rs110419, rs1027702, rs11037575, rs4336470, rs17065417) have been validated in an Italian population. These findings indicate that NB could arise from some as yet unknown combination of relatively common SNPs that can cumulatively increase the risk of a neuroblastic malignant transformation in fetal or early childhood development. No pairwise combinations of genotypes at the eight SNPs showed evidence of interaction, defined as deviation from additivity of allelic effects at separate loci. One of the major goals of this work was to determine the cumulative effect of these variants on NB risk. Performing separate genetic analyses in two independent populations, we have demonstrated that although individual susceptibility alleles only moderately increase the risk of NB, the risk becomes substantial when risk alleles are combined. When we tested for their cumulative effect, we found that children carrying 10–15 risk alleles had a 3.2-fold for European Americans and 5-fold for Italians increased risk of developing NB compared with those who carried 0–7 risk alleles, indicating the importance of the combined independent risk loci in neuroblast carcinogenesis. We observed similar results when the same analysis was performed on high-risk patients using the four genetic variants associated with clinically aggressive phenotypes (rs6939340, rs6435862, rs7585356 and rs110419). Indeed, European American children with NB carrying 6–8 risk alleles had 3.3-fold increased risk of developing high-risk NB compared with those who carried 0–3 risk alleles. In Italian patients the same risk was increased 1.6-fold, but it did not reach statistical significance.

One limitation of our study is the relative small number of Italian cases and controls. This probably affects some of the observed results. For instance, we speculate that the lack of association between some of SNPs and clinical conditions and the not statistically significant result of the risk allele cumulative analysis of high-risk patients are probably due to the limited power of the sample. However, in both analyses the trend toward association was similar in both Italian and European American populations. Surely, additional confirmatory studies are needed in larger populations with diverse origins to further validate the hypothesis that heritable DNA variation influences the clinical course of the disease.

Our study suggests that the heritable DNA variation that influences the initiation of NB and the clinical course of the disease has the potential to predict disease and more interestingly to predict high-risk NB. However, a large proportion of the heritability of NB remains undefined. Future studies will focus on discovering other common and rare disease susceptibility variants and epistatic effects at known risk alleles. We hypothesize that when a greater number of risk variants will be identified a more robust genetic score could be built, which may allow subgroups of patients at different risks of developing NB to be better distinguished. Several findings have demonstrated that NB is a type of tumor genetically and phenotypically heterogeneous (20). We think that the discriminatory power of combined risk alleles analysis will improve with time because the number of reproducibly associated variants is likely to increase rapidly with increasing number of cases and controls, as we have already demonstrated in our recent articles (5,7). We have planned to increase prominently the number of cases from three diverse populations (European Americans, Italians and British individuals) within the next 5 years through a multicentric collaborative project. A genetic risk score that includes more risk alleles could be relevant in improving the current risk models of NB adding more prediction power to the well-known clinical and genetic markers such as age, INSS stage, MYCN status. However, to assess the applicability of SNPs in the risk classification of NB, large prospective cohort studies will be needed.

In summary, our study confirms that the majority of risk loci identified in European American children also affect susceptibility to NB in Italian children and demonstrates compelling evidence of *BARD1* as the most significant genetic contributor to NB risk, exhibiting two independent risk variants. This work also shows that the identified NB

susceptibility loci individually have a moderate effect size but when combined may increase risk of NB substantially.

## Supplementary material

Supplementary Tables 1–5 and Figures 1–3 can be found at <http://carcin.oxfordjournals.org/>

## Funding

Associazione Italiana per la Ricerca sul Cancro (10537), MIUR- FIRB Ricerca in Futuro (RBFRO8DWQ3 to M.C.); Fondazione Italiana per la Lotta al Neuroblastoma (to M.C., F.C., L.L., M.D. and G.P.T.); Associazione Oncologia Pediatrica e Neuroblastoma (to M.C. and F.C.).

*Conflict of Interest Statement:* None declared.

## References

1. Maris, J.M. (2010) Recent advances in neuroblastoma. *N. Engl. J. Med.*, **362**, 2202–2211.
2. Cohn, S.L. *et al.* (2009) The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. *J. Clin. Oncol.*, **27**, 289–297.
3. Maris, J.M. *et al.* (2008) Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. *N. Engl. J. Med.*, **358**, 2585–2593.
4. Capasso, M. *et al.* (2009) Common variations in *BARD1* influence susceptibility to high-risk neuroblastoma. *Nat. Genet.*, **41**, 718–723.
5. Wang, K. *et al.* (2011) Integrative genomics identifies *LMO1* as a neuroblastoma oncogene. *Nature*, **469**, 216–220.
6. Nguyen, I.e.B. *et al.* (2011) Phenotype restricted genome-wide association study using a gene-centric approach identifies three low-risk neuroblastoma susceptibility Loci. *PLoS Genet.*, **7**, e1002026.
7. Diskin, S.J. *et al.* (2012) Identification of a new neuroblastoma susceptibility locus at 6q16 implicating *HACE1* and *LIN28B*. *Nat. Genet.*, **44**, 1126–1130.
8. Latorre, V. *et al.* (2012) Replication of neuroblastoma SNP association at the *BARD1* locus in African-Americans. *Cancer Epidemiol. Biomarkers Prev.*, **21**, 658–663.
9. Price, A.L. *et al.* (2008) Discerning the ancestry of European Americans in genetic association studies. *PLoS Genet.*, **4**, e236.
10. Seldin, M.F. *et al.* (2006) European population substructure: clustering of northern and southern populations. *PLoS Genet.*, **2**, e143.
11. Bosse, K.R. *et al.* (2012) Common variation at *BARD1* results in the expression of an oncogenic isoform that influences neuroblastoma susceptibility and oncogenicity. *Cancer Res.*, **72**, 2068–2078.
12. Holm, K. *et al.* (2010) SNPexp: a web tool for calculating and visualizing correlation between HapMap genotypes and gene expression levels. *BMC Bioinformatics*, **11**, 600.
13. Yang, T.P. *et al.* (2010) Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. *Bioinformatics*, **26**, 2474–2476.
14. Capasso, M. *et al.* (2007) Cytokine gene polymorphisms in Italian preterm infants: association between interleukin-10 -1082 G/A polymorphism and respiratory distress syndrome. *Pediatr. Res.*, **61**, 313–317.
15. Purcell, S. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
16. Johnson, A.D. *et al.* (2008) SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*, **24**, 2938–2939.
17. Barrett, J.C. *et al.* (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
18. Sporn, J.C. *et al.* (2011) *BARD1* expression predicts outcome in colon cancer. *Clin. Cancer Res.*, **17**, 5451–5462.
19. Irminger-Finger, I. *et al.* (2006) Is there more to *BARD1* than *BRCA1*? *Nat. Rev. Cancer*, **6**, 382–391.
20. Capasso, M. *et al.* (2010) Genetics and genomics of neuroblastoma. *Cancer Treat. Res.*, **155**, 65–84.

Received August 14, 2012; revised November 9, 2012; accepted November 30, 2012