

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Maris JM, Mosse YP, Bradfield JP, et al. Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. *N Engl J Med* 2008;358:2585-93. DOI: [10.1056/NEJMoa0708698](https://doi.org/10.1056/NEJMoa0708698).

SUPPLEMENTARY INFORMATION

Supplemental Methods

Genome-wide SNP Genotyping

DNA samples were surveyed for quality both by optical density spectrophotometry and the pico-green assay. Samples judged to be of sufficient quality for genotyping were assayed on the Illumina Infinium™ II HumanHap550 BeadChip technology,^{18, 19}(Illumina, San Diego) at the Center for Applied Genomics at the Children's Hospital of Philadelphia. A total of 750 nanograms of germline genomic DNA was used to genotype each sample, according to the manufacturer's guidelines. On day one, genomic DNA was amplified 1000-1500-fold; and on day two, the amplified DNA was fragmented to ~300-600 basepairs (bp), precipitated and resuspended followed by hybridization onto a BeadChip. Single base extension (SBE) utilized a single probe sequence of approximately 50 bp long designed to hybridize immediately adjacent to the single nucleotide polymorphism (SNP) query site. Following targeted hybridization to the bead array, the arrayed SNP locus-specific primers (attached to beads) were extended with a single hapten-labeled dideoxynucleotide in the single base extension reaction. The haptens were subsequently detected by a multi-layer immunohistochemical sandwich assay. The Illumina BeadArray Reader scanned each BeadChip at two wavelengths and created an image file. As BeadChip images were collected, intensity values were determined for all instances of each bead type, and data files were created that summarized intensity values for each bead type. These files consisted of intensity data that was loaded directly into Illumina's genotype analysis software, BeadStudio. A bead pool manifest created from the Laboratory Information Management System (LIMS) database containing all the BeadChip data was loaded into BeadStudio along with the intensity data for the samples. BeadStudio used a normalization algorithm to minimize BeadChip to BeadChip variability. Once the normalization was complete,

the clustering algorithm was run to evaluate cluster positions for each locus and assign individual genotypes. Each locus was given an overall score based on the quality of the clustering and each individual genotype call was given a GenCall score. GenCall scores provided a quality metric that ranges from 0 to 1 assigned to every genotype called. GenCall scores were then calculated using information from the clustering of the samples. The location of each genotype relative to its assigned cluster determined its GenCall score.

PCR-based genotyping for replication

All cases and controls from the UK and CCG/Los Angeles were genotyped at all SNPs showing a single marker association statistic in the discovery set below the genome-wide significance threshold by PCR-based allelic discrimination assays. Assays were performed at the respective sites according to the manufacturer's instructions using an ABI PRISM 7900HT system. Primers and probes were designed using Primer Express (Applied Biosystems) or Primer Picker (KBioscience) software (sequences available upon request). Assays were performed in 384-well format and each plate contained negative control wells. Data analysis was performed using Sequence Detection System software version 2.1 software (Applied Biosystems) and by visual inspection.

Table S1. Neuroblastoma patient and tumor characteristics of the 1032 cases used in the discovery set after exclusion of outliers in the Eigenstrat analysis of population stratification.

Characteristic	Number (%)
Sex	
Female	490 (48%)
Male	537 (52%)
Unknown	5
Age	
< 1 yr	386 (38%)
≥ 1 yr	638 (62%)
Unknown	8
INSS Stage	
Stage 1	217 (22%)
Stage 2	145 (14%)
Stage 3	156 (16%)
Stage 4	411 (42%)
Stage 4S	56 (6%)
Unknown	47
MYCN	
Not Amplified	817 (83%)
Amplified	167 (17%)
Not available	48
Histology	
Favorable	463 (57%)
Unfavorable	348 (48%)
Not available	221
DNA index	
Hyperdiploid	522 (67%)
Diploid	259 (33%)
Not available	251

Table S2. Neuroblastoma GWAS results in the discovery case series showing genotypes for all SNPs with genome-wide significant *P*-values. SNPs sorted by unadjusted *P*-value.

Chr	SNP	Genome Position (Build 36)	Gene	Ref Allele (A)	Minor Allele (B)	Cases				Controls				Allele test <i>P</i> -value
						AA	AB	BB	N	AA	AB	BB	N	
6	rs6939340	22247733	<i>FLJ22536</i>	A	G	198	504	310	1012	567	1015	451	2033	7.01x10 ⁻¹⁰
6	rs4712653	22233943	<i>FLJ22536</i>	T	C	208	503	285	996	604	1006	423	2033	1.16x10 ⁻⁹
6	rs9295536	22239908	<i>FLJ22536</i>	C	A	239	536	256	1031	667	1003	370	2040	1.71x10 ⁻⁹
20	rs3790171	19170336	<i>SLC24A3</i>	T	C	671	327	32	1030	1519	492	32	2043	3.64x10 ⁻⁸
20	rs7272481	19206143	<i>SLC24A3</i>	A	G	708	295	27	1030	1577	446	20	2043	4.99x10 ⁻⁸

Chr; chromosome. Genome position from NCBI build 36.

Table S3. Replication effort of chromosome 20 association results.

SNP	CHOP Case N	CHOP Case MAF	CHOP Control N	CHOP Control MAF	CHOP P- value	UK Case N	UK Case MAF	UK Control N	UK Control MAF	UK P- value	CCG Case N	CCG Case MAF	CCG Control N	CCG Control MAF	CCG P- value
rs3790171	394	0.1396	1178	0.1358	0.7897	246	0.1524	787	0.1823	0.322	ND	ND	ND	ND	ND
rs7272481	394	0.1244	1178	0.1231	0.9249	243	0.146	763	0.163	0.370	59	.1864	163	.3436	0.09.

Three separate patient groups selected to be of Northern European decent were used for replication. Cases and controls from CHOP were selected as for the discovery set (see methods). Neuroblastoma patients from the United Kingdom (UK) were identified from Pediatric Oncology Centers across the country, and controls were from the 1958 Birth Cohort Collection. Likewise, neuroblastoma patients from the Children's Cancer Group (CCG) Trial CCG-3891 were selected from the Neuroblastoma Resource Bank at Children's Hospital Los Angeles. ND, not done.

Table S4. Genotype frequency distributions in clinical subgroups and controls at three 6p22 SNPs.

	rs4712653					rs9295536					rs6939340				
	TT	CT	CC	OR (95%CI)	P	CC	AC	AA	OR (95%CI)	P	AA	AG	GG	OR (95%CI)	P
Stage 4	71 (17.8)	196 (49.1)	132 (33.1)	1.68* (1.16-2.44)	0.006	85 (20.7)	205 (50)	120 (29.3)	1.63* (1.13-2.34)	0.008	68 (16.9)	198 (49.2)	136 (33.8)	1.53* (1.05-2.22)	0.02
Not Stage 4	133 (23.4)	288 (50.7)	147 (25.9)	1.42 [†] (1.07-1.87)	0.015	150 (25.4)	311 (52.6)	130 (22.0)	1.47 [†] (1.09-1.96)	0.009	127 (21.8)	288 (49.6)	166 (28.6)	1.28 [†] (0.97-1.68)	0.08
MYCN Amp	26 (15.9)	70 (42.9)	67 (41.1)	2.17* (1.32-3.57)	0.002	32 (19.2)	74 (44.3)	61 (36.5)	2.04* (1.27-3.28)	0.003	25 (15.0)	74 (44.3)	68 (40.7)	1.99* (1.21-3.95)	0.006
MYCN Not Amp	173 (22.0)	408 (51.9)	205 (26.1)	1.98 [†] (1.39-2.81)	0.0001	196 (24.1)	437 (53.5)	183 (22.4)	1.99 [†] (1.39-2.84)	0.0001	166 (20.8)	406 (50.9)	226 (28.3)	1.74 [†] (1.23-2.45)	0.002
High-risk	65 (16.9)	191 (49.6)	129 (33.5)	1.83* (1.25-2.68)	0.002	77 (19.5)	198 (50.0)	121 (30.5)	1.95* (1.34-2.83)	0.0004	66 (16.9)	186 (47.7)	138 (35.4)	1.63* (1.11-2.38)	0.01
Not High-risk	130 (23.5)	281 (50.9)	141 (25.5)	1.47 [†] (1.10-1.95)	0.008	149 (25.9)	306 (53.2)	120 (20.9)	1.67 [†] (1.24-2.24)	0.0006	121 (21.5)	287 (51.0)	155 (27.5)	1.44 [†] (1.09-1.90)	0.01
Controls	604 (29.7)	1006 (49.5)	423 (20.8)			667 (32.7)	1003 (49.2)	370 (18.1)			567 (27.9)	1015 (49.9)	451 (22.2)		

*OR for risk allele homozygous genotype versus other homozygous genotype (CC vs TT; AA vs CC; GG vs AA) within each clinical category

[†]OR for risk allele homozygous genotype versus heterozygous + other homozygous genotypes (CC vs TT+CT; AA vs CC+AC; GG vs AA+AG) within each clinical category

Table S5: Allele frequency distributions in clinical subgroups and controls at three 6p22 SNPs.

	rs4712653				rs9295536				rs6939340			
	T (%)	C (%)	OR* (95% CI)	P*	C (%)	A (%)	OR* (95% CI)	P*	A (%)	G (%)	OR* (95% CI)	P*
Stage 4	338 (42.3)	460 (56.7)	1.63 (1.40-1.90)	3.92x10 ⁻¹⁰	375 (45.7)	445 (54.3)	1.59 (1.37-1.85)	1.28x10 ⁻⁹	334 (41.5)	470 (58.5)	1.58 (1.35-1.84)	4.05x10 ⁻⁹
Not Stage 4	554 (48.8)	582 (51.2)	1.26 (1.10-1.43)	0.0007	611 (51.7)	571 (48.3)	1.25 (1.10-1.43)	0.0006	542 (46.6)	620 (53.3)	1.28 (1.12-1.46)	0.0002
			1.29 (1.08-1.55)	0.005			1.27 (1.06-1.52)	0.009			1.23 (1.03-1.47)	0.02
MYCN Amp	122 (37.4)	204 (62.6)	2.00 (1.58-2.52)	2.89 x10 ⁻⁹	138 (41.3)	196 (58.7)	1.90 (1.52-2.39)	2 x10 ⁻⁸	124 (37.1)	210 (62.9)	1.89 (1.51-2.39)	3 x10 ⁻⁸
MYCN Not Amp	754 (48.0)	818 (52.0)	1.30 (1.15-1.46)	0.000012	829 (50.8)	803 (49.2)	1.30 (1.15-1.46)	0.000008	738 (46.3)	858 (53.7)	1.30 (1.16-1.46)	0.000007
			1.54 (1.20-1.97)	0.0005			1.46 (1.15-1.86)	0.002			1.45 (1.14-1.86)	0.002
High-risk	321 (41.7)	449 (58.3)	1.67 (1.43-1.95)	7.74x10 ⁻¹¹	352 (44.5)	440 (55.5)	1.68 (1.44-1.95)	3.91x10 ⁻¹¹	318 (40.8)	462 (59.2)	1.63 (1.39-1.90)	6.26 x10 ⁻¹⁰
Not High-risk	541 (49.0)	563 (51.0)	1.24 (1.08-1.42)	0.0013	604 (52.5)	546 (47.5)	1.21 (1.06-1.38)	0.004	529 (47.0)	597 (53.0)	1.26 (1.11-1.44)	0.0005
			1.34 (1.12-1.62)	0.002			1.38 (1.15-1.66)	0.0005			1.29 (1.07-1.55)	0.007
Controls	2214 (54.5)	1852 (45.5)			2337 (57.3)	1743 (42.7)			2149 (52.9)	1917 (47.1)		

*There are three odd-ratios (OR) and *P*-values listed for each SNP and clinical category. The first two are for allelic test of each subgroup versus controls and the third row is for the allelic test between the two subgroups.

Figure S1A. Quantile-quantile plot of genome-wide association results prior to Eigenstrat correction.

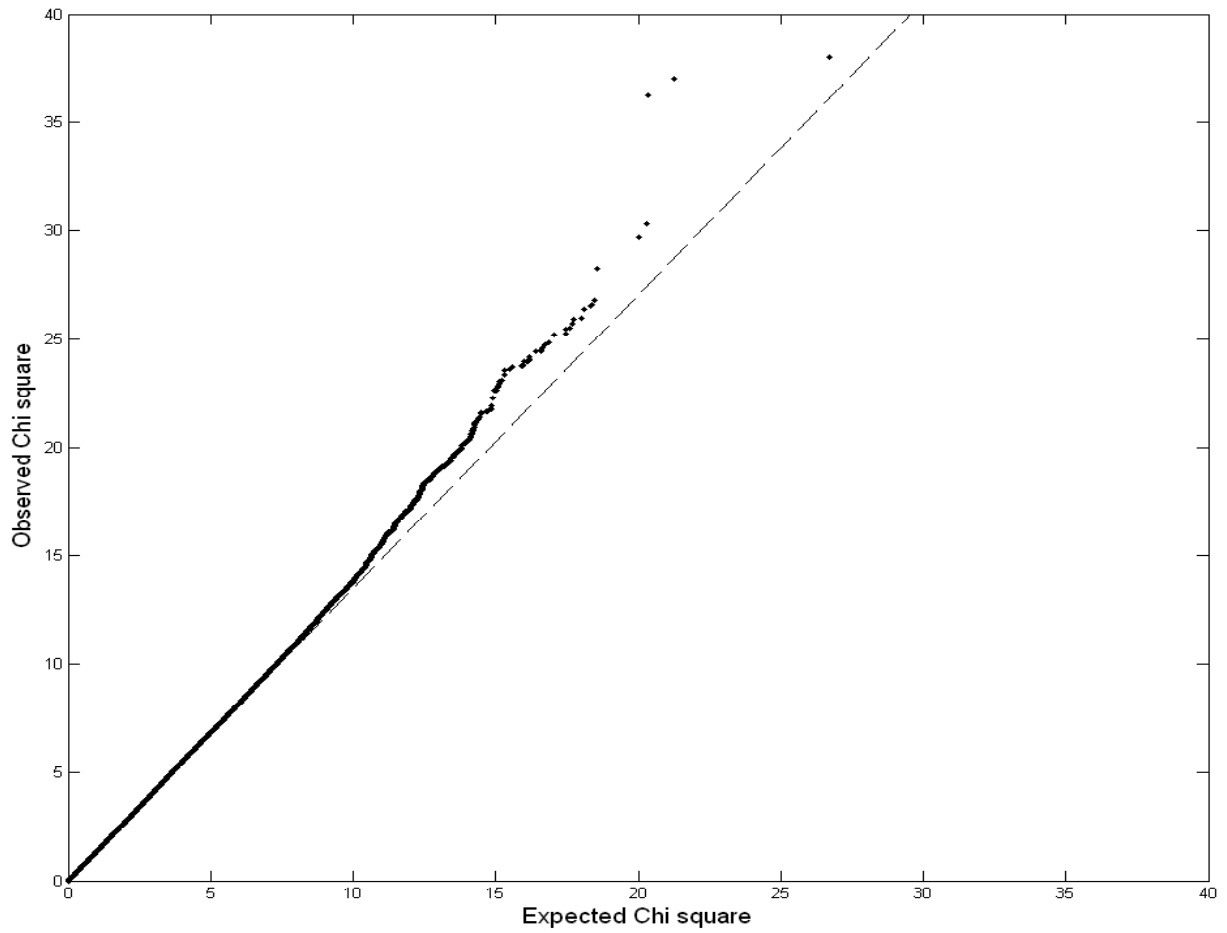


Figure S1B. Quantile-quantile plot of genome-wide association results after Eigenstrat correction showing that many of the potentially genome-wide significant SNPs in S1A are no longer significant, and the only three remaining genome-wide significant SNPs are ones mapping in tight linkage disequilibrium at 6p22.

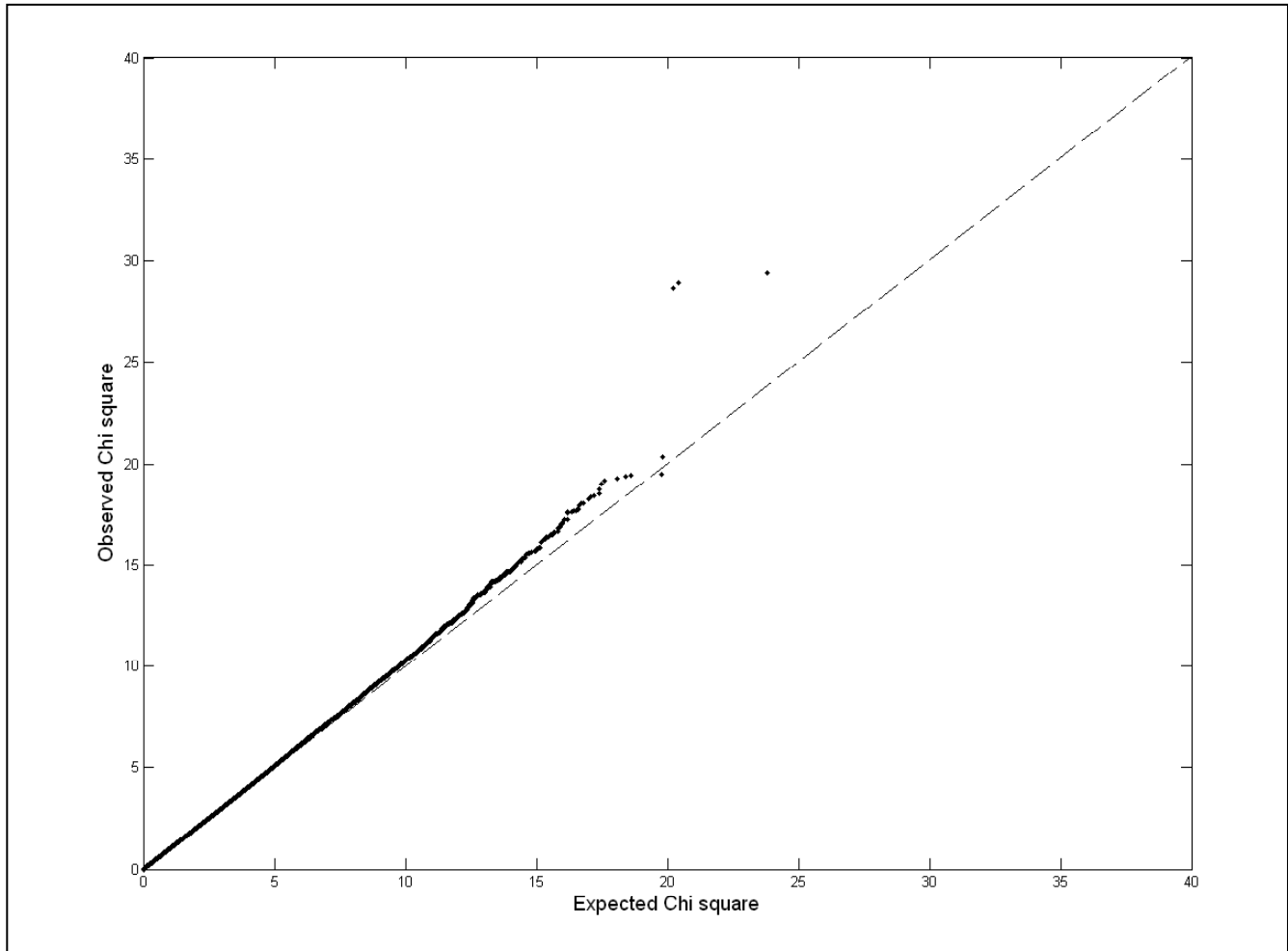


Figure S2. Event-free survival probabilities for neuroblastoma patients with the homozygous at-risk genotype compared to those homozygous for the non-risk allele at rs6939340.

