



Published in final edited form as:

N Engl J Med. 2008 June 12; 358(24): 2585–2593. doi:10.1056/NEJMoa0708698.

A genome-wide association study identifies a susceptibility locus to clinically aggressive neuroblastoma at 6p22

John M. Maris^{1,2}, Yael P. Mosse^{1,*}, Jonathan P. Bradfield^{3,*}, Cuiping Hou¹, Stefano Monni⁴, Richard H. Scott⁵, Shahab Asgharzadeh⁶, Edward F. Attiyeh¹, Sharon J. Diskin¹, Marci Laudenslager¹, Cynthia Winter¹, Kristina Cole¹, Joseph T. Glessner³, Cecilia Kim³, Edward C. Frackelton³, Tracy Casalunovo³, Andrew W. Eckert³, Mario Capasso^{7,8}, Eric F. Rappaport¹, Carmel McConville⁹, Wendy B. London¹⁰, Robert C. Seeger⁶, Nazneen Rahman⁵, Marcella Devoto^{4,8,11}, Struan F. A. Grant^{3,8}, Hongzhe Li⁴, and Hakon Hakonarson^{3,8}

¹Division of Oncology and Center for Childhood Cancer Research, Children's Hospital of Philadelphia; Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

²Abramson Family Cancer Research Institute, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

³The Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

⁴Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

⁵Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK

⁶Division of Hematology — Oncology and Saban Research Institute, Children's Hospital Los Angeles, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

⁷University "Federico II", CEINGE - Biotechnologie Avanzate, Naples, Italy

⁸Division of Genetics, The Children's Hospital of Philadelphia; Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

⁹Division of Reproductive & Child Health and CRUK Institute for Cancer Studies, University of Birmingham, UK

¹⁰University of Florida and Children's Oncology Group, Department of Statistics, Gainesville, FL, USA

¹¹University of Rome "La Sapienza", Department of Experimental Medicine, Rome, Italy

Abstract

Background—Neuroblastoma is a malignancy of the developing sympathetic nervous system that most commonly affects young children and is often lethal. The etiology of this embryonal cancer is not known.

Correspondence: John M. Maris, MD Children's Hospital of Philadelphia ARC902A 3516 Civic Center Boulevard Philadelphia, PA 19104-4318 Tel: 215 590 2821 Fax: 215 590 3770 Email: maris@chop.edu Hakon Hakonarson, MD, PhD Children's Hospital of Philadelphia ARC12 3516 Civic Center Boulevard Philadelphia, PA 19104-4318 Tel: 267 426 0088 Fax: 267 426 0363 Email: hakonarson@chop.edu.

*These authors contributed equally. The authors declare no competing financial interests.

DISCLOSURE

The authors have no financial disclosures. Dr. Maris authored the original manuscript draft.

Methods—We performed a genome-wide association study by first genotyping 1,032 neuroblastoma patients and 2,043 controls of European descent using the Illumina HumanHap550 BeadChip. Three independent groups of neuroblastoma cases (N=720) and controls (N=2128) were then genotyped to replicate significant associations.

Results—We observed highly significant association between neuroblastoma and the common minor alleles of three single nucleotide polymorphisms (SNPs) within a 94.2 kilobase (Kb) linkage disequilibrium block at chromosome band 6p22 containing the predicted genes *FLJ22536* and *FLJ44180* (P -value range = 1.71×10^{-9} - 7.01×10^{-10} ; allelic odds ratio range 1.39-1.40). Homozygosity for the at-risk G allele of the most significantly associated SNP, rs6939340, resulted in an increased likelihood of developing neuroblastoma of 1.97 (95% CI 1.58-2.44). Subsequent genotyping of these 6p22 SNPs in the three independent case series confirmed our observation of association ($P=9.33 \times 10^{-15}$ at rs6939340 for joint analysis). Furthermore, neuroblastoma patients homozygous for the risk alleles at 6p22 were more likely to develop metastatic (Stage 4) disease ($P=0.02$), show amplification of the *MYCN* oncogene in the tumor cells ($P=0.006$), and to have disease relapse ($P=0.01$).

Conclusion—Common genetic variation at chromosome band 6p22 is associated with susceptibility to neuroblastoma.

Despite marked improvements in the cure rates for many childhood cancers, neuroblastoma remains an important clinical problem accounting for 15% of the mortality attributable to malignancy in children.¹ It is the most common solid cancer of early childhood, and approximately half of all neuroblastoma cases present with widely disseminated disease that is often refractory to intensive chemoradiotherapy. Cure rates for these high-risk patients remain less than 40% despite dramatic increases in chemoradiotherapy intensity, and survivors often suffer serious life-long morbidities.²⁻⁴ Somatically acquired genomic aberrations in neuroblastoma are of fundamental importance for predicting tumor phenotype in patients.¹ Tumors with amplification of the *MYCN* oncogene, or deletions of chromosome arms 1p and/or 11q, typically are metastatic at diagnosis and resistant to therapy.⁵ Conversely, tumors showing no structural chromosomal changes, but hyperdiploidy due to whole chromosome gains, are more easily cured and may even spontaneously regress.⁶ Despite the wealth of knowledge on somatically acquired genomic aberrations that correlate with tumor phenotype, little is known about the genetic events that predispose to neuroblastoma tumorigenesis.

Knudson and Strong originally predicted that neuroblastoma fits the two mutation model of oncogenesis,⁷ subsequently proven for the analogous embryonal cancer retinoblastoma.⁸ However, a family history of neuroblastoma is obtained in only about 1% of cases,⁹ and to date genetic studies in these rare families show locus heterogeneity with no commonly mutated gene identified.¹⁰⁻¹² Mutations in the retinoblastoma gene (*RBI*) confer several orders of magnitude increase in the risk of developing tumor, penetrance is nearly complete, and mutations in this suppressor gene are involved in essentially all familial and sporadic cases; whereas neuroblastoma appears to be more genetically complex and heterogeneous. We therefore hypothesized that the majority of neuroblastomas arise from relatively common DNA variations that predispose to an increased risk for neuroblastic malignant transformation.

METHODS

Research Subjects

For genome-wide genotyping, cases were defined as a child diagnosed with neuroblastoma or ganglioneuroblastoma and registered through the Children's Oncology Group (COG). The blood samples from the neuroblastoma cases were identified through the COG Neuroblastoma bio-repository for specimen collection at the time of diagnosis. The majority of specimens were annotated with clinical and genomic information that included: age at diagnosis, site of origin,

disease stage by the International Neuroblastoma Staging System¹³, INPC International Neuroblastoma Pathology Classification¹⁴, *MYCN* oncogene copy number,¹⁵ DNA index (ploidy),¹⁶ registration on clinical trial(s), event-free and overall survival, second malignancies, and any associated conditions (e.g. congenital abnormalities).

Eligibility criterion for genome-wide genotyping was availability of 1.5 µg of high quality DNA from a tumor-free source such as peripheral blood or uninvolved (with tumor) bone marrow mononuclear cells. Because neuroblastoma in the United States is demographically a disease of Caucasians of European descent,¹⁷ we limited our initial analyses to this racial group to minimize genetic variability. The neuroblastoma cases genotyped genome-wide were randomly divided into a discovery set of 1251 cases, and an initial replication set of 409 cases, all annotated as self-reported Caucasian.

Control subjects were recruited from the Philadelphia region through the Children's Hospital of Philadelphia (CHOP) 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: 1) self-reported as Caucasian; 2) availability of 1.5 µg of high quality DNA from peripheral blood mononuclear cells; and 3) no serious underlying medical disorder including cancer. A total of 3414 control subjects genotyped genome-wide were used in this study, randomly allocated to 2236 used in the discovery phase, and 1178 in the initial replication effort. The median age of the control subjects at the time of sample collection was 10.0 years.

Two additional case series were used for the purpose of replication. First, 252 randomly selected and unrelated individuals with neuroblastoma or ganglioneuroblastoma were recruited from Pediatric Oncology Centers in the United Kingdom either as part of a collection of sporadic neuroblastoma patients diagnosed at the Birmingham Children's Hospital since 1992, or through the Factors Associated with Childhood Tumours (FACT) study or from the Childhood Cancer and Leukaemia Group Tumour and Leukaemia Bank. Control samples (N=788) were from the 1958 Birth Cohort Collection, an ongoing follow-up study of all persons born in Great Britain during one week in 1958, including a recent biomedical assessment during 2002–2004 at which blood samples and informed consent were obtained for creation of a genetic resource (National Child Development Study <http://www.cls.ioe.ac.uk/>). All cases and controls were from the United Kingdom and subjects known to be of non-European ethnic groups were excluded. The final replication group of 59 unrelated individuals with high-risk neuroblastoma were recruited from the US-based Children's Cancer Group (CCG) protocols from the 1990's². Control DNA samples for the CCG cases (N=162) were collected from unrelated Caucasian individuals living in the Los Angeles area during 2000–2001 from buccal swabs obtained for the creation of an anonymized genetic resource.

Informed consent was obtained for all participants and this study was approved by each participating center's Institutional Review Board as well as the COG Scientific Council, COG Neuroblastoma Disease Committee and Cancer Therapy Evaluation Program (CTEP) at the NCI.

Genotyping

Details of methods for genome-wide genotyping have been described^{18,19}, and are included with methods for replication genotyping by PCR-based allelic discrimination assays in the Supplementary Information.

Data analysis and statistical methods

Genome-wide genotyping data from an initial 1,251 neuroblastoma patient and 2,236 disease-free control discover set were filtered based on pre-specified quality control measures where

individual SNPs were excluded from further analysis if they showed: 1) deviation from Hardy Weinberg equilibrium with $P < 0.0001$; 2) individual SNP genotype yield $< 98\%$; or 3) minor allele frequency (MAF) $< 5\%$. This resulted in 464,934 SNPs being utilized in the subsequent analyses. In addition, 33 samples had genotype yields $< 90\%$ and were removed (23 cases and 10 controls). Finally, since the case samples were accrued nation-wide, while the control set was recruited locally in Philadelphia, we performed principal components analyses to identify outlier samples in order to reduce the effects of population stratification.^{20,21} This approach removed 379 samples (196 cases and 183 controls), leaving 1,032 cases and 2,043 controls for our discovery case series with self-reported ethnicity of Caucasian. Evaluation of these 3,075 subjects using ancestry informative markers (AIMs) available on the HumanHap550 BeadChip predicted Caucasian status in all but two subjects (these two subjects remained in the analysis).²³ The neuroblastoma subjects were representative of the expected distribution of clinical and biological covariates as observed in the general population (Supplemental Table S1).^{1,17}

The primary statistical tests for association in the discovery case series were carried out using the software package PLINK.²² We conservatively set 1.0×10^{-7} as the threshold for genome-wide significance, based on the fact that slightly less than 500,000 SNPs were used in the analysis ($0.05/500,000 = 1.0 \times 10^{-7}$). The single marker analyses for the genome-wide data were carried out using the χ^2 test based on allele count differences between 1,032 cases and 2,043 controls and the Cochran-Armitage test for trends on genotype frequencies. Allelic odds ratios and the corresponding 95% confidence intervals were calculated for the association analyses. In addition, to further control for the potential confounding influence of population stratification, we performed association analyses after correction for substructure based on principal components analysis as implemented in Eigenstrat^{20,21}. For each SNP, we utilized the default settings within the program and performed a modified Cochran-Armitage trend test adjusting for the top five principal components²¹, and report the result as the Eigenstrat *p*-value. Supplementary Figure S1 shows quantile-quantile plots before and after correction.

Of the 1,032 patients included in the discovery case series, clinical and biological covariate data obtained at diagnosis was available for most (Supplementary Table S1). A total of 883 patients (85.6%) had complete outcome data available with a median follow-up interval of 4.02 years for patients without an event. Association analyses of chromosome 6p22 SNPs with clinical characteristics were performed with the χ^2 test on allele and genotype counts, and with outcome by comparing Kaplan-Meier survival curves by the logrank test in a pairwise fashion.

The Eigenstrat corrected summary statistics for the entire discovery case series are available at <http://www.caglab.net/pub/kiran/neuroblastoma.html>.

RESULTS

To identify sequence variants that are associated with susceptibility to develop neuroblastoma, we compared single marker allele and genotype frequencies in our discovery case series using the χ^2 and Cochran-Armitage trend test statistics. The top three SNPs showing significant association to neuroblastoma were in tight linkage disequilibrium (LD) at chromosome 6p22 (rs6939340, rs4712653, rs9295536) yielding *P*-values of 1.71×10^{-9} - 7.01×10^{-10} (allelic odds ratio 1.39-1.40; Figure 1, Table 1, Table S2). Two additional SNPs at chromosome 20p11 (rs3790171 and rs7272481) showed genome-wide significant single marker *P*-values, and many others were very close to the genome-wide significance threshold (Figure 1). However, only the chromosome 6 association results retained genome-wide significance after further correction for population substructure using principal components analyses as implemented in Eigenstrat^{20,21} (Table 1). The chromosome 6 signal falls within a 94.2 Kb LD block containing the predicted overlapping genes *FLJ22536* and *FLJ44180* (Figure 2).

We next sought to replicate the chromosome 6p22 and 20p11 association signals in three separate pairings of neuroblastoma patients with controls. As shown in Table 2, the chromosome 6 risk alleles identified in the discovery phase were also significantly over-represented in all three sets of neuroblastoma cases compared to their controls, yielding a combined *P*-value and allelic odds ratio for the most strongly associated SNP rs6939340 of 9.33×10^{-15} and 1.37 (95% CI 1.27-1.49). There was no support for the chromosome 20 alleles identified in the discovery cohort truly being associated with predisposition to neuroblastoma in the three replication case series (Supplementary Table S3).

To determine if the risk alleles at the chromosome 6p22 locus were differentially associated with patient subsets and outcome, we analyzed the associated SNPs frequency distribution from the discovery phase COG case series with respect to prognostically relevant clinical and biological covariates present at diagnosis and survival rates. For all three SNPs, there was a significantly different allele frequency distribution indicating that the risk alleles were more likely to be present in patients with a more malignant clinical presentation and disease course (Figure 3). Accordingly, neuroblastoma patients homozygous for the at-risk alleles were more likely to develop tumors that were metastatic at diagnosis (*P*=0.0060, 0.0084, and 0.0245 for rs4712653, rs9295536, and rs6939340 respectively), to have somatically acquired amplification of the *MYCN* oncogene in tumor cells (*P*=0.0018, 0.0028, and 0.0061 respectively), and to have a high-risk classification for the purposes of treatment stratification (*P*=0.0018, 0.0004 and 0.0107 respectively), compared to neuroblastoma patients homozygous for the non-risk alleles (Supplemental Table S4 and S5). In addition, patients homozygous for the at-risk alleles had a significantly decreased event-free survival probability (*P*=0.0158 for SNP rs6939340; Supplemental Figures S2). This association with more malignant disease may explain why we were able to show replication of results in the relatively small series of 59 cases from CCG that consisted solely of high-risk patients (Table 2).

DISCUSSION

Embryonal cancers are postulated to arise from partially committed primordial cells during fetal or early childhood development. This comprehensive genetic analysis shows that the likelihood for malignant transformation of developing neuroblasts is influenced by common variation in the human genome at 6p22, and that the ultimate neuroblastoma phenotype is in part determined by germline variation at this locus. These definitive data support a model of pediatric cancers as complex disease, with common variations in the genome likely influencing susceptibility and phenotypic variability.

The motivation for this large and ongoing study was the paucity of information on how neuroblastoma tumorigenesis is initiated. Epidemiological studies have not identified a common environmental exposure that influences susceptibility to neuroblastoma,²⁴⁻²⁶ and genetic studies of hereditary disease have been hampered by the rarity of the condition and the small size of pedigrees due to the lethality of neuroblastoma in early childhood.²⁷ Our first results not only provide proof-of-concept for a genome-wide association approach to neuroblastoma, but also identifies novel candidates for susceptibility genes. Little is known about the predicted genes within the LD block identified here. *FLJ22536* has multiple predicted isoforms and contains a potential epidermal growth factor-like domain. The *FLJ44180* gene is apparently novel as there are no sequence similarities in human or mouse nucleic acids and protein databases. Studies are underway to determine how the presence of common DNA variants at 6p22 contributes to the risk of neuroblastic malignant transformation, presumably through altered expression and/or alternative splicing of regional candidate genes. The association of common SNP alleles at 6p22 with the sub-phenotype of high-risk disease also requires further exploration. Future work will define why tumors initiated with a genetic

contribution from the 6p22 locus are more likely to evolve towards a more malignant clinical disease.

Figure 1 demonstrates that there are multiple other potential association signals that will emerge as we continue towards our eventual goal of genotyping 5000 neuroblastoma cases genome-wide. While speculative, these early data support a common variant-common disease model for neuroblastoma where initiation of neuroblastoma tumorigenesis occurs due to the interaction of multiple relatively common genetic variations in the developing neuroblast. Our ongoing genome-wide association study and international collaborations are expected to identify and verify additional risk alleles and how these may interact to increase susceptibility to neuroblastoma.

Neuroblastoma is a unique malignancy which includes the highest proportion of cases that will spontaneously regress compared to any other human cancer,^{28,29} but the majority of patients have a relentless malignant disease that is difficult to cure even with intensive chemoradiotherapy.¹ Our data suggest that the chromosome 6p22 risk alleles not only influence the likelihood of developing neuroblastoma, but also the likelihood for developing a more malignant phenotype. This observation increases the urgency for understanding the molecular pathways influenced by these risk variants in somatic and tumor tissues, as this may lead to novel approaches to neuroblastoma therapy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was made possible due to the efforts of multiple investigators throughout the Children's Oncology Group (U10-CA98543, Dr. Gregory Reaman Chair). This work was also supported in part by NIH Grants R01-CA78454 (JMM), R01-ES009911 (HL) and R01-CA60104 (RCS), the Alex's Lemonade Stand Foundation (JMM), the Center for Applied Genomics at the Joseph Stokes Research Institute, which funded all GWA genotyping (HH), and the Abramson Family Cancer Research Institute (JMM), the Neuroblastoma Society (NR, RHS), the Institute of Cancer Research, UK (NR, RHS). We acknowledge use of DNA from the British 1958 Birth Cohort collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. MC is supported in part by a fellowship from the Associazione Italiana Ricerca sul Cancro (AIRC). The Factors Associated with Childhood Tumours (FACT) Study and the Tumour and Leukemia Bank and Children's Cancer and Leukaemia Group (CCLG) also supported these studies.

REFERENCES

1. Maris JM, Hogarty MD, Bagatell R, Cohn SL. Neuroblastoma. *Lancet* 2007;369(9579):2106–20. [PubMed: 17586306]
2. Matthay KK, Villablanca JG, Seeger RC, et al. Treatment of high-risk neuroblastoma with intensive chemotherapy, radiotherapy, autologous bone marrow transplantation, and 13-cis-retinoic acid. Children's Cancer Group. *New England Journal of Medicine* 1999;341(16):1165–73. [PubMed: 10519894]
3. Berthold F, Boos J, Burdach S, et al. Myeloablative megatherapy with autologous stem-cell rescue versus oral maintenance chemotherapy as consolidation treatment in patients with high-risk neuroblastoma: a randomised controlled trial. *Lancet Oncol* 2005;6(9):649–58. [PubMed: 16129365]
4. Oeffinger KC, Mertens AC, Sklar CA, et al. Chronic health conditions in adult survivors of childhood cancer. *N Engl J Med* 2006;355(15):1572–82. [PubMed: 17035650]
5. Attiyeh EF, London WB, Mosse YP, et al. Chromosome 1p and 11q deletions and outcome in neuroblastoma. *N Engl J Med* 2005;353(21):2243–53. [PubMed: 16306521]
6. George RE, London WB, Cohn SL, et al. Hyperdiploidy plus nonamplified MYCN confers a favorable prognosis in children 12 to 18 months old with disseminated neuroblastoma: a Pediatric Oncology Group study. *J Clin Oncol* 2005;23(27):6466–73. [PubMed: 16116152]

7. Knudson AGJ, Strong LC. Mutation and cancer: Neuroblastoma and pheochromocytoma. *American Journal of Human Genetics* 1972;24:514–22. [PubMed: 4340974]
8. Friend SH, Bernards R, Rogelj S, et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 1986;323(6089):643–6. [PubMed: 2877398]
9. Shojaei-Brosseau T, Chompret A, Abel A, et al. Genetic epidemiology of neuroblastoma: A study of 426 cases at the Institut Gustave-Roussy in France. *Pediatric Blood and Cancer* 2004;42(1):99–105. [PubMed: 14752801]
10. Maris JM, Weiss MJ, Mosse Y, et al. Evidence for a hereditary neuroblastoma predisposition locus at chromosome 16p12-13. *Cancer Res* 2002;62(22):6651–8. [PubMed: 12438263]
11. Trochet D, Bourdeaut F, Janoueix-Lerosey I, et al. Germline Mutations of the Paired-Like Homeobox 2B (PHOX2B) Gene in Neuroblastoma. *Am J Hum Genet* 2004;74(4):761–4. [PubMed: 15024693]
12. Longo L, Panza E, Schena F, et al. Genetic predisposition to familial neuroblastoma: identification of two novel genomic regions at 2p and 12p. *Hum Hered* 2007;63(34):205–11. [PubMed: 17317969]
13. Brodeur GM, Pritchard J, Berthold F, et al. Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *Journal of Clinical Oncology* 1993;11(8):1466–77. [PubMed: 8336186]
14. Shimada H, Ambros IM, Dehner LP, Hata J, Joshi VV, Roald B. The International Neuroblastoma Pathology Classification (Shimada) System. *Cancer* 1999;86:364–72. [PubMed: 10421273]
15. Mathew P, Valentine MB, Bowman LC, et al. Detection of MYCN gene amplification in neuroblastoma by fluorescence in situ hybridization: a pediatric oncology group study. *Neoplasia* 2001;3(2):105–9. [PubMed: 11420745]
16. Look AT, Hayes FA, Shuster JJ, Douglass EC, Castleberry RP, Brodeur GM. Clinical relevance of tumor cell ploidy and N-myc gene amplification in childhood neuroblastoma. A Pediatric Oncology Group Study. *Journal of Clinical Oncology* 1991;9:581–91. [PubMed: 2066755]
17. Brodeur, GM.; Maris, JM. Neuroblastoma. In: Pizzo, PA.; Poplack, DG., editors. *Principles and Practice of Pediatric Oncology*. Vol. 5th. J. B. Lippincott Company; Philadelphia: 2006. p. 933-70.
18. Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat Genet* 2005;37(5):549–54. [PubMed: 15838508]
19. Steemers FJ, Chang W, Lee G, Barker DL, Shen R, Gunderson KL. Whole-genome genotyping with the single-base extension assay. *Nat Methods* 2006;3(1):31–3. [PubMed: 16369550]
20. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38(8):904–9. [PubMed: 16862161]
21. Patterson N, Price AL, Reich D. Population Structure and Eigenanalysis. *PLoS Genet* 2006;2(12):e190. [PubMed: 17194218]
22. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81(3):559–75. [PubMed: 17701901]
23. Yang N, Li H, Criswell LA, et al. Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. *Hum Genet* 2005;118(34):382–92. [PubMed: 16193326]
24. Olshan AF, De Roos AJ, Teschke K, et al. Neuroblastoma and parental occupation. *Cancer Causes and Control* 1999;10:539–49. [PubMed: 10616823]
25. De Roos AJ, Olshan AF, Teschke K, et al. Parental occupational exposures to chemicals and incidence of neuroblastoma in offspring. *Am J Epidemiol* 2001;154(2):106–14. [PubMed: 11447042]
26. De Roos AJ, Teschke K, Savitz DA, et al. Parental occupational exposures to electromagnetic fields and radiation and the incidence of neuroblastoma in offspring. *Epidemiology* 2001;12(5):508–17. [PubMed: 11505168]
27. Maris, JM.; Brodeur, GM. Genetics. In: Cheung, N-KV.; Cohn, SL., editors. *Neuroblastoma*. Springer; Berlin, Heidelberg, New York: 2005. p. 21-6.
28. Everson, TC.; Cole, WH. Spontaneous regression of neuroblastoma. In: Everson, TC.; Cole, WH., editors. *Spontaneous Regression of Cancer*. W. B. Saunders Co.; Philadelphia: 1966. p. 88-163.

29. Carlsen NL. How frequent is spontaneous remission of neuroblastomas? Implications for screening. *Br J Cancer* 1990;61(3):441–6. [PubMed: 2328213]
30. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21(2):263–5. [PubMed: 15297300]

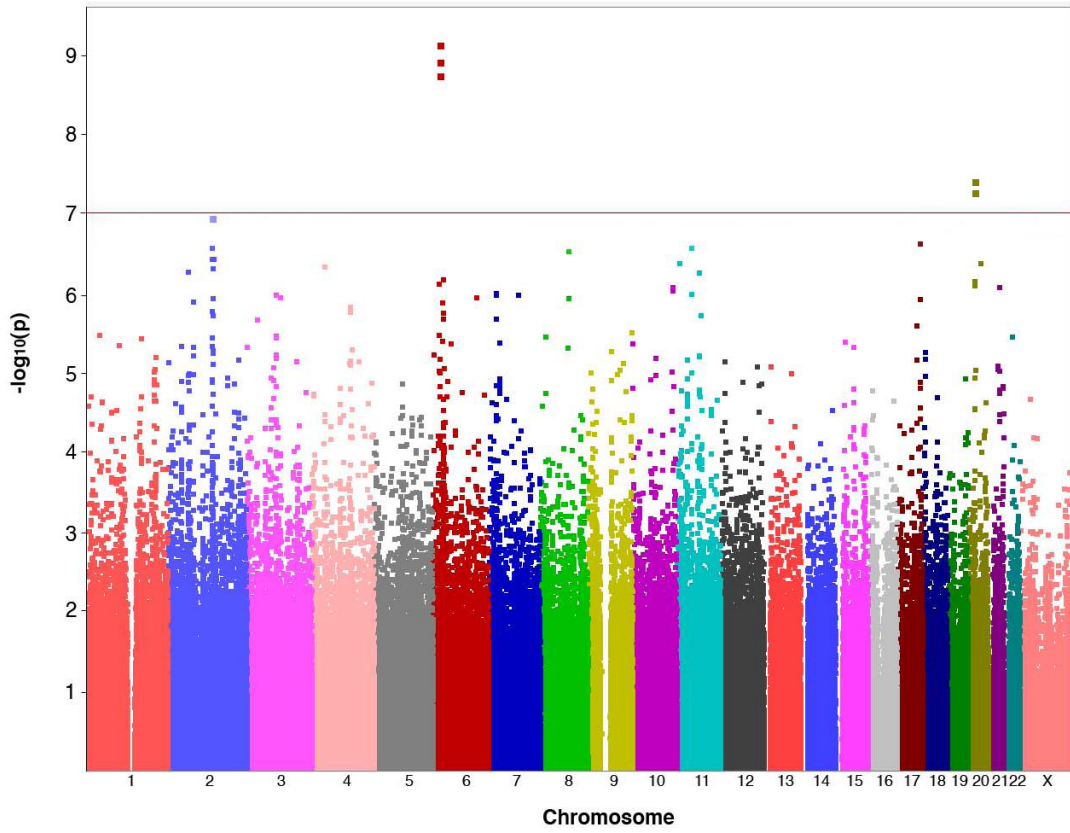


Figure 1. Summary of the neuroblastoma GWAS results in the discovery set
The y-axis represents the level of significance for each SNP (log transformed P -values) at the relative genomic position on each chromosome along the x-axis from short arm terminus (left) to long arm terminus (right). The red line indicates the genome-wide significance threshold.

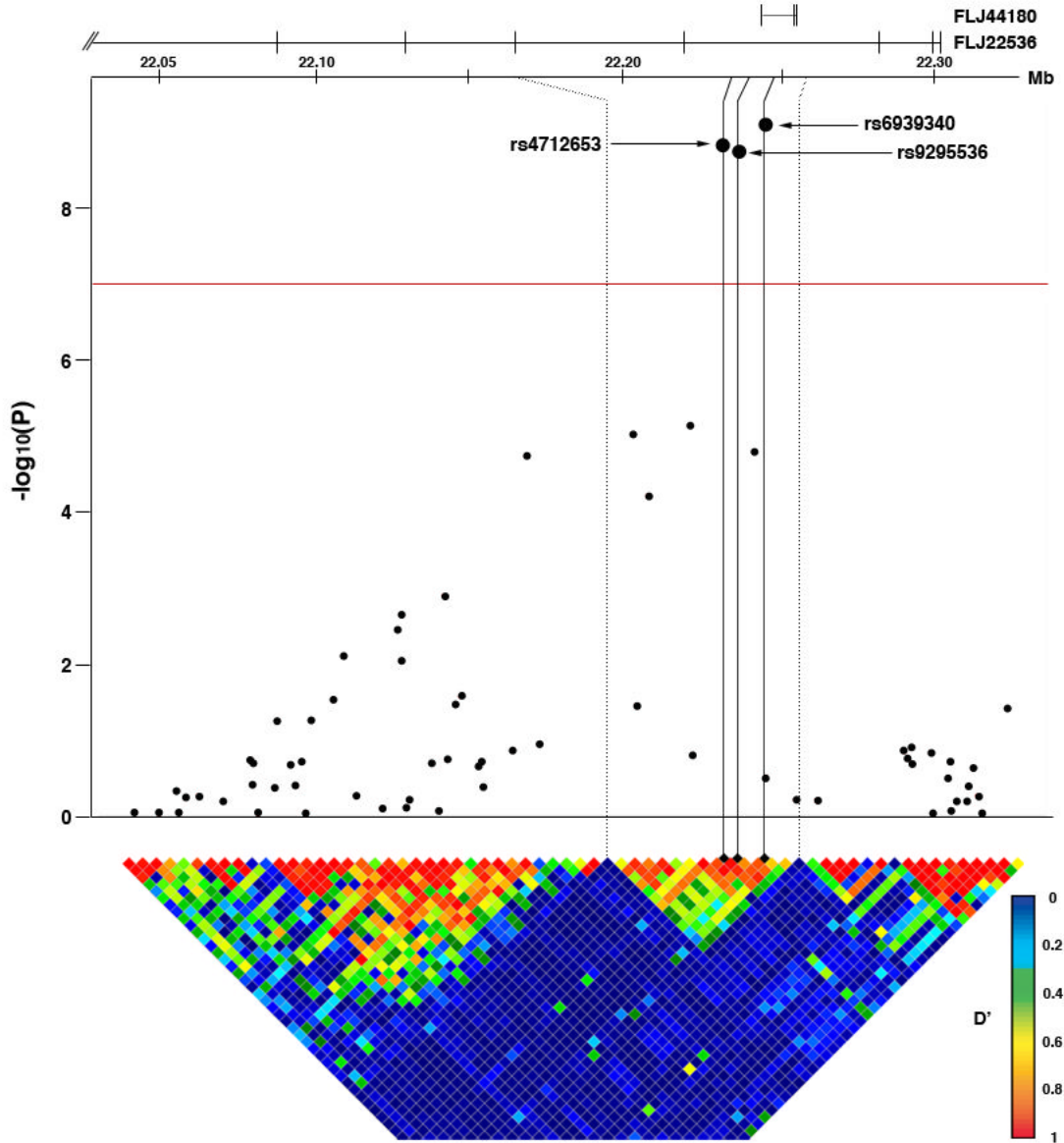


Figure 2. Pairwise linkage disequilibrium diagram of the 6p22 locus and association results for regional SNPs

A 290.98 Kb region from rs12207699 to rs9393266 is shown, with pairwise LD calculated from the discovery case and control data set combined (LD measured as D' as indicated by scale). Individual SNP log transformed P -values are plotted in relation to the LD diagram, both created using Haploview.³⁰ The three SNPs showing a genome-wide significant association signal (red horizontal line) are annotated (r^2 was 0.873 for SNPs rs4712653 and rs9295536; and 0.731 for SNPs rs9295536 and rs6939340) and map within a 94.2 Kb LD block, delimited by SNPs rs196051 and rs10498705 (dashed vertical lines). All SNPs are also annotated to their relative physical map positions (NCBI Build 36), with predicted genes in the region shown (exons/transcribed sequences indicated by dashed vertical lines).

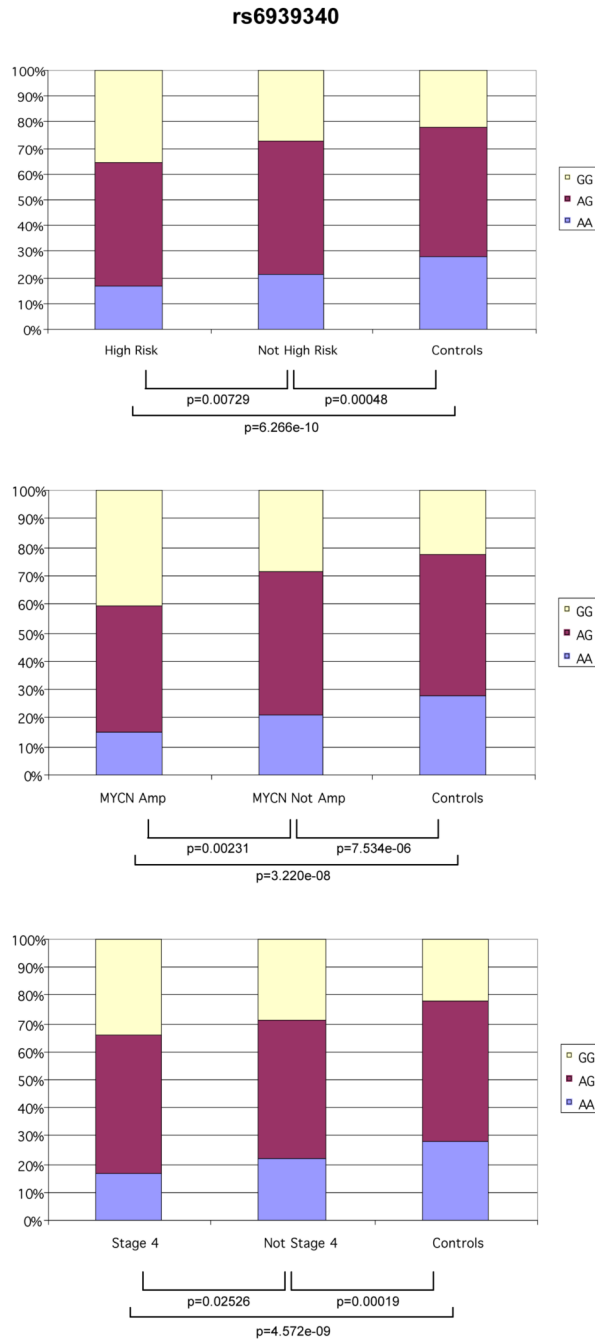


Figure 3. Chromosome 6p22 risk alleles are associated with a more malignant neuroblastoma phenotype

Genotype frequency distributions (homozygous G, heterozygous AG, or homozygous A) of patients classified according to three clinical variables and controls are shown for SNP rs6939340, along with p-values of allelic tests between the clinical subgroups, and between each of them and the controls. Similar results are obtained for SNPs rs4712653 and rs9295536. Specifically, patients assigned to a high-risk group, with tumors that harbor *MYCN* amplification, and with Stage 4 disease were more likely to be homozygous for the risk allele C at rs4712653, A at rs9295536 and G at rs6939340 (see text and Supplementary Tables S4 and S5).

Table 1
Neuroblastoma genome-wide association study results for discovery case and control subjects.
 SNPs yielding genome-wide significant association, sorted by unadjusted *P*-value.

Chr	SNP	Genome Position (Build 36)	Gene	Ref Allele (A)	Minor Allele (B)	Case N	Case MAF	Control N	Control MAF	χ^2 P-value	Cochran-Armitage Trend P-Value	Eigenstrat P-value	OR (95% CI)	[AB- AA] OR (95% CI)	[BB- AA] OR (95% CI)
6	rs6939340	22247733	<i>FLJ22536</i>	A	G	1012	0.5553	2033	0.4715	7.01×10^{-10}	7.33×10^{-10}	5.82×10^{-8}	1.40 (1.26-1.56)	1.42 (1.17-1.73)	1.97 (1.58-2.45)
6	rs4712653	22233943	<i>FLJ22536</i>	T	C	996	0.5387	2033	0.4555	1.16×10^{-9}	1.21×10^{-10}	7.43×10^{-8}	1.40 (1.25-1.55)	1.45 (1.20-1.76)	1.96 (1.57-2.43)
6	rs9295536	22239908	<i>FLJ22536</i>	C	A	1031	0.5082	2040	0.4272	1.71×10^{-9}	1.40×10^{-9}	8.54×10^{-8}	1.39 (1.25-1.54)	1.49 (1.24-1.79)	1.93 (1.55-2.40)
20	rs3790171	19170336	<i>SLC24A3</i>	T	C	1030	0.1898	2043	0.1361	3.64×10^{-8}	2.54×10^{-8}	8.85×10^{-5}	1.49 (1.29-1.71)	1.50 (1.27-1.78)	2.26 (1.38-3.73)
20	rs7272481	19206143	<i>SLC24A3</i>	A	G	1030	0.1694	2043	0.1189	4.99×10^{-8}	3.26×10^{-8}	6.96×10^{-5}	1.51 (1.30-1.75)	1.47 (1.24-1.75)	3.01 (1.68-5.40)

Statistical tests for association were carried out using the software package PLINK (<http://pngu.mgh.harvard.edu/purcell/plink/>)²². The single marker analysis for the genome-wide data was carried out using a χ^2 test on allele count differences between cases and controls. Results of the Cochran-Armitage test for trends and Eigenstrat analysis are also included (see text). Allelic odds ratios (OR) and the corresponding 95% confidence intervals were calculated for the association analysis and are shown. Odds ratio are shown for the allelic test, as well as for heterozygosity [AB-AA] and homozygosity [BB-AA] for the B risk allele. Chr: chromosome. MAF: minor allele frequency. Genome position from NCBI build 36 (hg 18).

Table 2

Replication effort of chromosome 6 association results.

SNP	CHOP Case N	CHOP Case MAF	CHOP Control N	CHOP Control MAF	CHOP P-value	CHOP OR (95% CI)	UK Case N	UK Case MAF	UK Control N	UK Control MAF	UK P-value	UK OR (95% CI)	CCG Case N	CCG Case MAF	CCG Control N	CCG Control MAF	CCG P-value	CCG OR (95% CI)	Combined P-value	Combined OR (95% CI)
rs6939340	394	0.5546	1178	0.4958	0.0043	1.27 (1.08-1.49)	214	0.5678	785	0.4771	0.0009	1.44 (1.16-1.79)	59	0.6356	161	0.4969	0.0098	1.77 (1.12-2.80)	9.33×10⁻¹⁵	1.37 (1.27-1.49)
rs4712653	394	0.5431	1178	0.4873	0.0066	1.25 (1.07-1.47)	205	0.5220	766	0.4576	0.0219	1.29 (1.04-1.61)	59	0.6102	161	0.4845	0.0194	1.67 (1.06-2.63)	5.50×10⁻¹³	1.35 (1.24-1.46)
rs9295536	394	0.5152	1178	0.4660	0.0167	1.22 (1.04-1.43)	244	0.4959	782	0.4380	0.0248	1.26 (1.03-1.55)	59	0.5488	162	0.4512	0.0635	1.49 (0.98-2.89)	1.24×10⁻¹¹	1.32 (1.22-1.43)

Three separate patient groups selected to be of Northern European descent were used for replication. Cases and controls from CHOP were selected as for the discovery case series (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. Control DNAs were derived from buccal swabs of individuals participating in a Los Angeles-based collection of adult control subjects for an anonymized genetic resource. The final two columns show the P-values the odds ratio (OR) and 95% confidence interval for the discovery and replication sets combined. MAF: minor allele frequency.