



Data in Brief

DNA methylation fingerprint of neuroblastoma reveals new biological and clinical insights



Soledad Gómez^a, Giancarlo Castellano^b, Gemma Mayol^a, Ana Queiros^b, José I. Martín-Subero^{b,c}, Cinzia Lavarino^{a,*}

^a Developmental Tumor Biology Laboratory, Hospital Sant Joan de Déu, Fundació Sant Joan de Déu, Barcelona, Spain

^b Institut d' Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain

^c Department of Anatomic Pathology, Pharmacology and Microbiology, University of Barcelona, Barcelona, Spain

ARTICLE INFO

Article history:

Received 9 July 2015

Accepted 13 July 2015

Available online 17 July 2015

Keywords:

Neuroblastoma
Embryonal tumor
High-density microarray
DNA methylation
HM450K

ABSTRACT

Neuroblastoma (NB) is one of the most frequently occurring extracranial solid tumors of childhood (Maris et al., 2007 [1]; Brodeur, 2003 [2]). Probability of cure varies according to patient's age, extent of disease and tumor biology (Maris et al., 2007 [1]; Brodeur, 2003 [2]; Cohn et al., 2009 [3]). However, the etiology of this developmental tumor is unknown. Recent evidence has shown that pediatric solid tumors, including NB, harbor a paucity of recurrent genetic mutations, with a significant proportion of recurrent events converging on epigenetic mechanisms (Cheung et al., 2012 [4]; Molenaar et al., 2012 [5]; Pugh et al., 2013 [6]; Sausen et al., 2013 [7]).

We have analyzed the DNA methylome of neuroblastoma using high-density microarrays (Infinium Human Methylation 450k BeadChip) to define the epigenetic landscape of this pediatric tumor and its potential clinicopathological impact. Here, we provide the detail of methods and quality control parameters of the microarray data used for the study. Methylation data has been deposited at NCBI Gene Expression Omnibus data repository, accession number GSE54719; superseries record GSE54721.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Specifications

Organism/cell line/tissue	<i>Homo sapiens</i>
Sex	Neuroblastoma samples (n = 35) were composed by 17 males and 18 females. All reference samples (n = 4) were males.
Sequencer or array type	Infinium Human Methylation 450k BeadChip (Illumina Inc., San Diego, CA). GPL13534
Data format	Raw data is available as TXT non-normalized methylated and non-normalized unmethylated tables.
Experimental factors	Tumor and reference samples.
Experimental features	Normal fetal brain (n = 2) and adrenal gland (n = 2) tissues as well as ganglioneuroma (benign Schwannian stroma dominant neuroblastic tumor) samples (n = 2) were used as reference samples. Tumor cell content was evaluated by a pathologist.
Consent	Patients, parents or guardians signed an informed consent before sample collection.
Sample source location	Hospital Sant Joan de Déu – Barcelona, Passeig Sant Joan de Déu, 2, 08950 Esplugues de Llobregat, Barcelona, Spain. Latitude: 41.383906 longitude: 2.101771

1. Direct link to deposited data

Methylation microarray data has been deposited at NCBI Gene Expression Omnibus data repository (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) with accession number GSE54719 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54719>); on superseries record GSE54721 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54721>).

2. Experimental design

To define the DNA methylation landscape of neuroblastoma and its potential clinicopathological impact. Microarray DNA methylation data (Infinium Human Methylation 450k BeadChip) were analyzed and associated with functional/regulatory genome annotation data, transcriptional profiles and clinico-biological parameters.

3. Material and methods

3.1. Patients and samples

Thirty-five primary neuroblastoma tumors (including 6 stage 1, 9 stage 3, 6 stage 4s and 14 stage 4) obtained at the time of diagnosis

* Corresponding author.

E-mail address: clavarino@fsjd.org (C. Lavarino).

Table 1
Patients' clinical and biological characteristics.

Characteristics	Methylation array samples (n = 35)		
Age, months	Median	17.90	
	Range	0.10–225.28	
INSS, n (%)	Stages 1–3	15	44.12%
	Stage 4	14	41.18%
	Stage 4s	6	17.65%
MYCN status, n (%)	MYCN Amp	7	20.59%
	MYCN non-Amp	28	82.35%

from patients diagnosed and treated at Hospital Sant Joan de Déu (Barcelona, Spain) were included in the study. Normal fetal brain (n = 2) and adrenal gland (n = 2) tissues as well as ganglioneuroma samples (n = 2) were used as reference samples. Neuroblastoma risk assessment was defined by the International Neuroblastoma Staging System (INSS) [8]. Patients' clinical and biological characteristics are displayed in Table 1. Tumors were evaluated by a pathologist, only samples with more than 70% viable tumor cell content were included in the study. This study was approved by the Institutional Review Boards. Informed consent was obtained before collection of samples.

Genomic DNA was isolated from samples using Wizard DNA Purification Kit (Promega Biotech Ibérica, Spain), following manufactures' protocols.

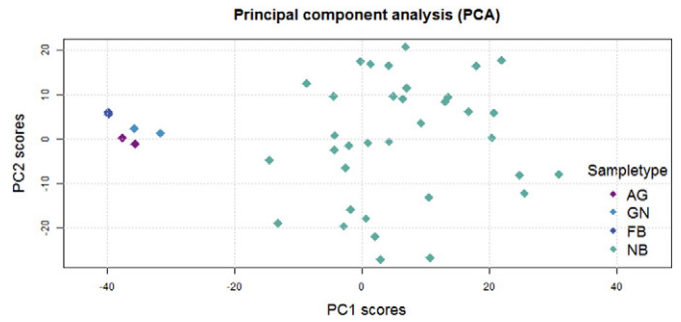


Fig. 2. Principal component analysis. AG: adrenal gland. GN: ganglioneuroma. FB: fetal brain. NB: neuroblastoma.

3.2. DNA methylation microarrays

Genomic DNA bisulfite conversion and hybridization to Infinium Human Methylation 450k BeadChip (Illumina Inc., San Diego, CA) was performed at the Human Genotyping Unit of the Spanish National Cancer Center (CEGEN-CNIO, Madrid, Spain), as described previously [9].

3.3. Quality control

Microarray data was analyzed by minfi [10] and shinyMethyl [11] packages available through Bioconductor [12].

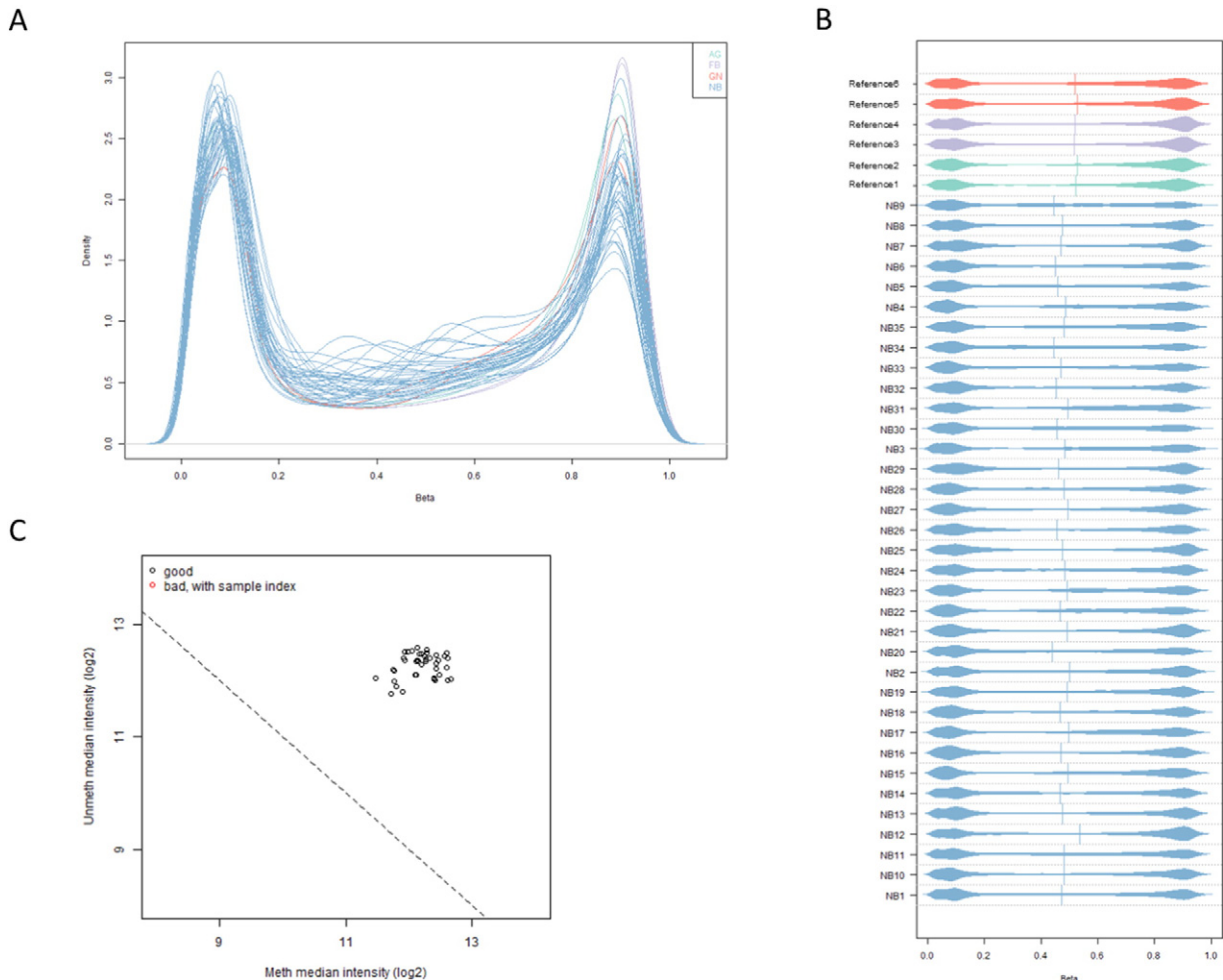


Fig. 1. A: Raw β -value density plot. B: Raw β -value bean density plot. C: Scatter plot of log median intensity of methylated (Meth) and unmethylated (Unmeth) channels.

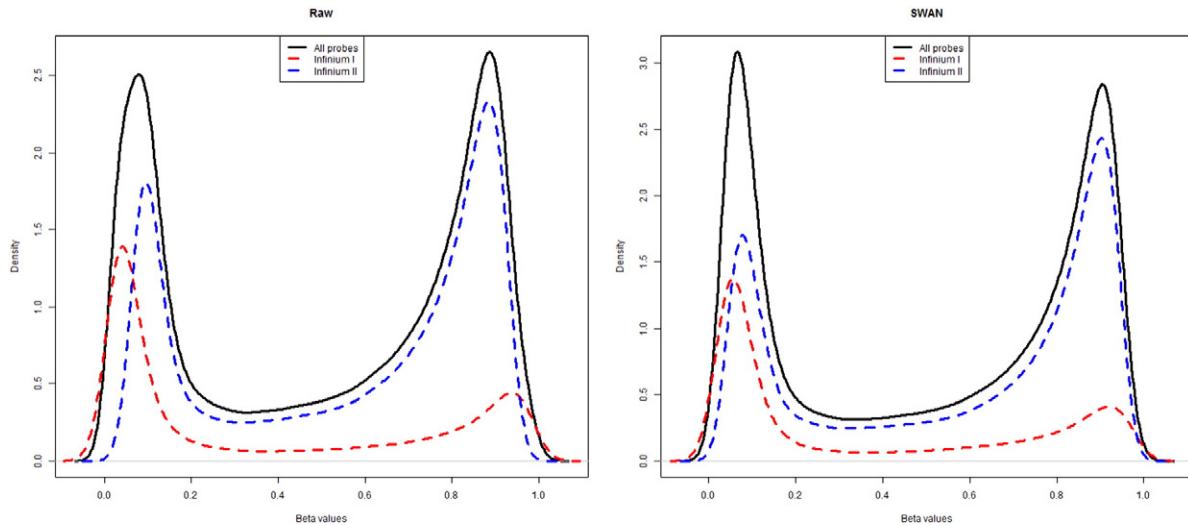


Fig. 3. Density distributions of beta values before and after using SWAN.

Single cytosine methylation values (β -values) in each sample were calculated as the ratio of the methylated signal intensity to the sum of methylated and unmethylated signals. Methylation values ranged from 0, fully unmethylated, to 1, fully methylated cytosine. Raw methylation β -value plots were performed for all samples to identify possible deviant samples (Fig. 1 & B).

Sample quality was assessed using the log median intensity in both the methylated and unmethylated channels (Fig. 1C).

Principal component analysis (PCA) was performed on the 20,000 most variable autosomal probes in order to analyze associations between sample type (adrenal gland, ganglioneuroma, fetal brain or neuroblastoma) and methylation levels (Fig. 2).

3.4. Data normalization

Microarray methylation data was normalized by missMethyl package [13] using the SWAN (subset-quantile within array normalization), a within-array normalization method for Illumina Human Methylation 450k BeadChips (Fig. 3).

3.5. Filtering of microarray data

An optimized pipeline with several filters developed at the IDIBAPS was used to exclude technical biases, [14,15]. Cytosines with poor detection P-values ($P > 0.01$) and those with sex-specific DNA methylation ($n = 6926$) were removed from the initial dataset of 485,512 sites (excluding probes detecting SNPs). The remaining sites ($n = 478,169$) were used for downstream analyses (Fig. 4).

3.6. Bisulfite pyrosequencing

DNA methylation data were validated by bisulfite pyrosequencing as previously described [15,16] (Fig. 5).

4. Data summary

The analysis of the DNA methylome of neuroblastoma patients revealed that: i) most DNA methylation changes take place outside promoters, ii) not only CpGs but also non-CpG sites are targeted, iii) methylation changes at these sites are associated with clinicopathological features of neuroblastoma, and iv) the identified epigenetic patterns provide new insights into the pathogenesis and clinical behavior of this pediatric tumor. The analysis of this data has been reported in our recently published study [15].

Financial disclosure

This work was supported by the NEN association (association of families and friends of patients with neuroblastoma) [to S Gómez], Hospital Sant Joan de Déu, Barcelona, Spain [BR201102 to G. Mayol] and Spanish Government [PI11/02661 and PI14/00038 Instituto de Salud Carlos III grant to C Lavarino and Ramon y Cajal Grant to JI Martín-Subero].

Acknowledgments

The authors acknowledge G. Garcia-Castellví for the fund raising and the “Biobanc de l’Hospital Infantil Sant Joan de Déu per a la Investigació” integrated in the National Network Biobanks of ISCIII for tissue samples. The authors also thank J. Ríos for statistical advice and S. Sánchez Molina for precious technical support.

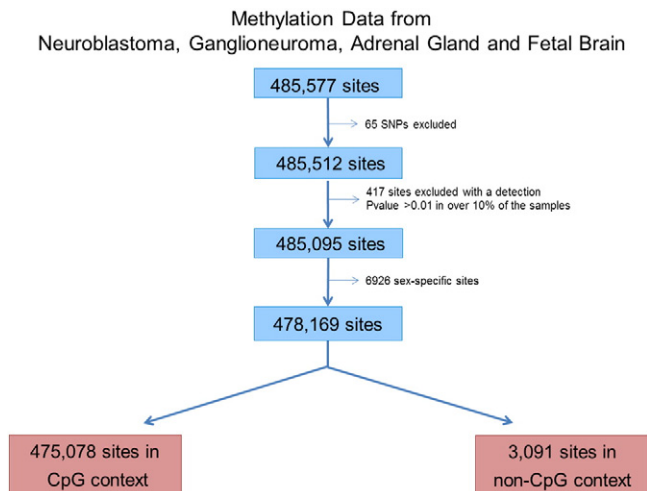


Fig. 4. Data processing work-flow.

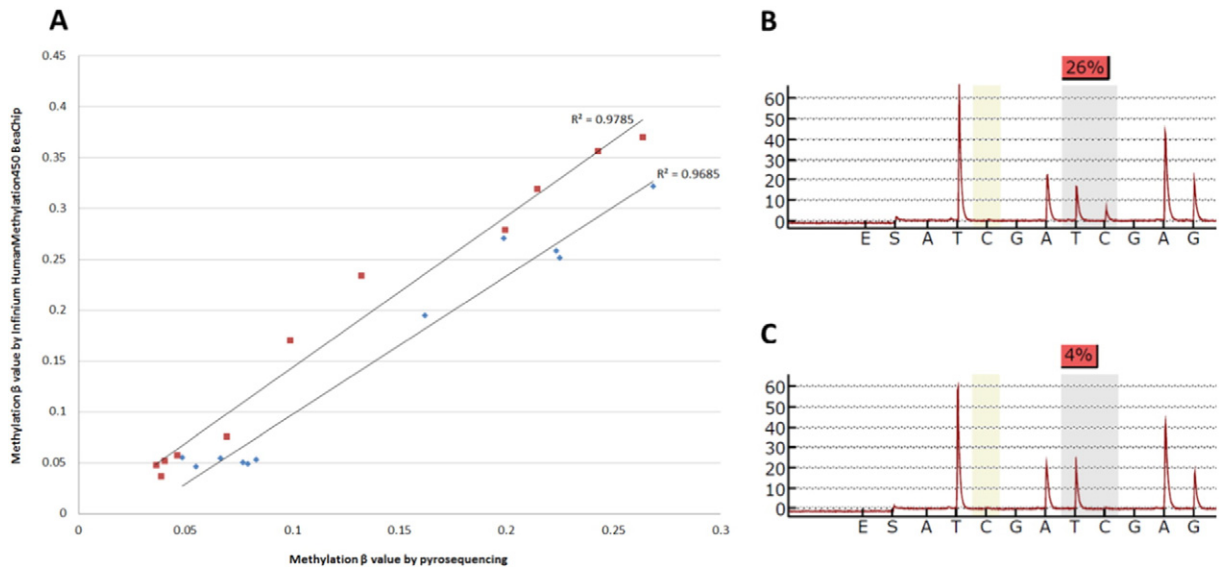


Fig. 5. Technical DNA methylation validation. (A) Scatter plot of DNA methylation levels obtained by Infinium Human Methylation 450k BeadChip (Illumina Inc., San Diego, CA) (Y-axis) and bisulfite pyrosequencing (X-axis) ($n = 11$). A highly significant correlation ($P < 0.001$, $R^2 = 0.978$ and $R^2 = 0.968$, respectively) between microarray and pyrosequencing data was observed. Example of pyrograms with methylated (B) and unmethylated (C) cytosines.

References

- [1] J.M. Maris, M.D. Hogarty, R. Bagatell, S.L. Cohn, Neuroblastoma. *Lancet* 369 (9579) (2007) 2106–2120.
- [2] G.M. Brodeur, Neuroblastoma: biological insights into a clinical enigma. *Nat. Rev. Cancer* 3 (3) (2003) 203–216.
- [3] S.L. Cohn, A.D. Pearson, W.B. London, et al., The International Neuroblastoma Risk Group (INRG) classification system: an INRG task force report. *J. Clin. Oncol.* 27 (2) (2009) 289–297.
- [4] N.K. Cheung, J. Zhang, C. Lu, et al., Association of age at diagnosis and genetic mutations in patients with neuroblastoma. *JAMA* 307 (10) (2012) 1062–1071.
- [5] J.J. Molenaar, J. Koster, D.A. Zwijnenburg, et al., Sequencing of neuroblastoma identifies chromothripsis and defects in neurogenesis genes. *Nature* 483 (7391) (2012) 589–593.
- [6] T.J. Pugh, O. Morozova, E.F. Attiyeh, et al., The genetic landscape of high-risk neuroblastoma. *Nat. Genet.* 45 (3) (2013) 279–284.
- [7] M. Sausen, R.J. Leary, S. Jones, et al., Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. *Nat. Genet.* 45 (1) (2013) 12–17.
- [8] G.M. Brodeur, J. Pritchard, F. Berthold, et al., Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 11 (8) (1993) 1466–1477.
- [9] M. Bibikova, B. Barnes, C. Tsan, et al., High density DNA methylation array with single CpG site resolution. *Genomics* 98 (4) (2011) 288–295.
- [10] M.J. Aryee, A.E. Jaffe, H. Corrada-Bravo, et al., Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 30 (10) (2014) 1363–1369.
- [11] J.P. Fortin, E. Fertig, K. Hansen, shinyMethyl: interactive quality control of Illumina 450 k DNA methylation arrays in R. *F1000Research* 3 (2014) 175.
- [12] W. Huber, V.J. Carey, R. Gentleman, et al., Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* 12 (2) (2015) 115–121.
- [13] J. Maksimovic, L. Gordon, A. Oshlack, SWAN: subset-quantile within array normalization for Illumina Infinium HumanMethylation450 BeadChips. *Genome Biol.* 13 (6) (2012) R44.
- [14] M. Kulis, S. Heath, M. Bibikova, et al., Epigenomic analysis detects widespread genome-wide DNA hypomethylation in chronic lymphocytic leukemia. *Nat. Genet.* 44 (11) (2012) 1236–1242.
- [15] S. Gomez, G. Castellano, G. Mayol, et al., DNA methylation fingerprint of neuroblastoma reveals new biological and clinical insights. *Epigenomics* (2015) 1–17.
- [16] G. Mayol, J.I. Martin-Subero, J. Rios, et al., DNA hypomethylation affects cancer-related biological functions and genes relevant in neuroblastoma pathogenesis. *PLoS One* 7 (11) (2012) e48401.